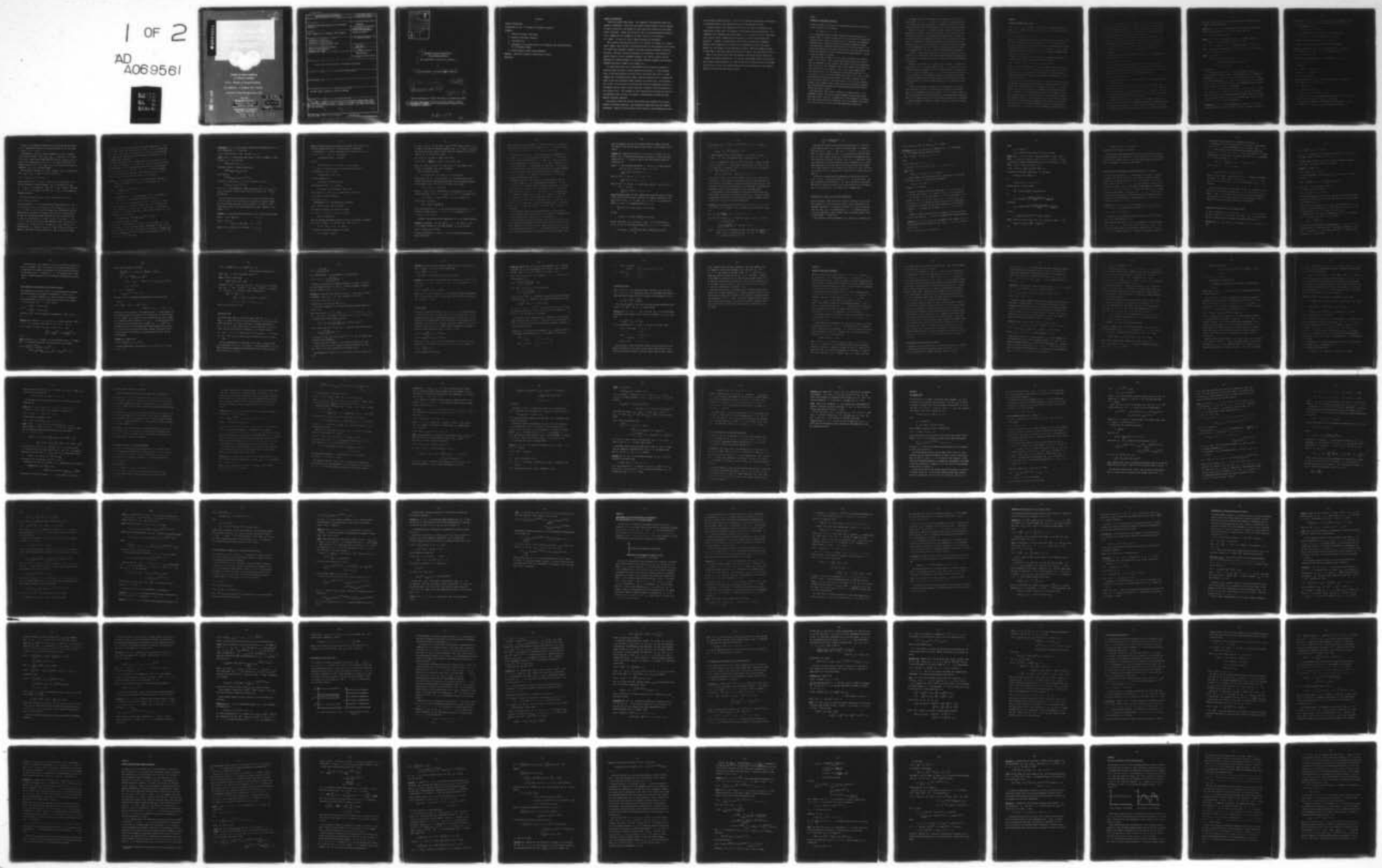
AD-A069 561 NORTH CAROLINA UNIV AT CHAPEL HILL INST OF STATISTICS F/G 12/1
EXTREMAL AND RELATED PROPERTIES OF STATIONARY PROCESSES. PART I--ETC(U)
MAY 79 M R LEADBETTER, G LINDGREN, H ROOTZEN N00014-75-C-0809
UNCLASSIFIED MIMEO-SER-1227 NL

1 OF 2
AD
A069561



B.S. 7

LEVEL II

AD A069561

[redacted]

EXTREMAL AND RELATED PROPERTIES

OF STATIONARY PROCESSES

Part I: Extremes of Stationary Sequences

M.R. Leadbetter, G. Lindgren, and H. Rootzén

Institute of Statistics Mimeo Series #1227

May 1979

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited



DEPARTMENT OF STATISTICS
Chapel Hill, North Carolina

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

## REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

Extremal and Related Properties of Stationary Processes

5. TYPE OF REPORT & PERIOD COVERED

TECHNICAL

6. PERFORMING ORG. REPORT NUMBER

Mimeo Series No. 1227

7. AUTHOR(s)

M.R. Leadbetter, G. Lindgren, and H. Rootzen

8. CONTRACT OR GRANT NUMBER(s)

N00014-75-C-0809

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina 27514

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS

Statistics & Probability Program
Office of Naval Research
Arlington, VA 22217

12. REPORT DATE

May 1979

13. NUMBER OF PAGES

95

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)

UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release: Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Extreme values, stochastic processes, maxima

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This resport considers the generalization of classical extreme value theory for independent random variables, to apply to stationary stochastic processes. Part 1 is concerned with stochastic sequences and part 2 will deal with continuous time processes.



DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

6

**EXTREMAL AND RELATED PROPERTIES**
**OF STATIONARY PROCESSES.**

**Part I. Extremes of stationary sequences,**

by

10

M.R. Leadbetter, G. Lindgren, ~~and~~ H. Rootzén

9 Technical rept.,

12 103p.

14 MIMEO-SER-1227

11 May 79

American Mathematical Society 1970 subject classification: 60G10.

15 Research supported by The Office of Naval Research, Contract ~~No.~~ N00014-75-C-0809, and The Swedish Natural Research Council, Contract No. F3741-004.

410 064

LB

Contents

## General Introduction

Classical Extreme Value Theory - the asymptotic distributional theory for maxima of independent, identically distributed random variables, may be regarded as half a century old, even though its roots reach much further back into mathematical antiquity. During this period of time it has found significant application - exemplified best perhaps by the book "Statistics of Extremes" by E.J. Gumbel - as well as a rather complete theoretical development.

More recently, beginning with work of G.S. Watson, S.M. Berman, R.M. Loynes, and H. Cramér, there has been a developing interest in the extension of the theory to include first dependent sequences, and then continuous parameter stationary processes. The early activity proceeded in two directions - the extension of general theory to certain dependent sequences (e.g. Watson, Loynes) and the beginning of a detailed theory for stationary sequences (Berman) and continuous parameter processes (Cramér) in the normal case.

In recent years both lines of development have been actively pursued, but especially that relating to normal sequences and processes. In the sequence case, it has proved possible to unify the two directions and to give a rather complete and satisfying general theory along the classical lines, including the known results for stationary normal sequences as special cases. Our principal aim in Part I of this work is to present this theory for dependent sequences in as complete and up to date a form as possible, alongside a brief description of the classical case. The treatment is thus unified both as regards the classical and dependent cases, and also in respect of consideration of normal and more general stationary sequences.

Some general theory has recently been obtained for extremes of *continuous* parameter stationary processes, and illuminating comparisons with the sequence case emerge. However the most detailed results known in the continuous case are

for stationary normal processes. In Part II we therefore concentrate substantially on the normal theory, and consider the general situation more briefly.

Closely related to the properties of extremes are those of exceedances and upcrossings of high levels, by sequences and continuous parameter processes. By regarding such exceedances and upcrossings as *point processes*, one may obtain some quite general results demonstrating convergence to Poisson and related point processes. A number of interesting results follow concerning the asymptotic behavior of the magnitude and location of such quantities as the k-th largest maxima (or local maxima, in the continuous setting). These and a number of other related topics, have been taken up in both Part I and II, but especially in Part II.

Many of the results given here have appeared in print in various forms, but a number are hitherto unpublished. For the previously known results we have given the smoothest proof (sometimes also new) of which we are aware and our aim throughout has been to stress the underlying main principles which pervade (and connect) both the classical and the present context.

## PART I

## EXTREMES OF STATIONARY SEQUENCES

Classical extreme value theory is concerned substantially with distributional properties of the maximum

$$M_n = \max(\xi_1, \xi_2, \ldots, \xi_n)$$

of $n$ independent and identically distributed random variables, as $n$ becomes large. Our primary object in Part I of this work is to extend the classical theory to apply to certain *dependent* sequences. The sequences considered form a large subclass of the stationary sequences – namely those exhibiting a dependence structure which is "not too strong", in a sense to be made precise. We shall find, in fact, that the classical theory may be extended in a rather complete and satisfying manner, to apply to this more general situation.

Our first task – in Chapter 1 – is to give an account of the relevant parts of the classical theory, emphasizing those results with which we shall be concerned in later chapters. Two results are of basic importance here. The first is the fundamental result which exhibits the possible limiting forms for the distribution of $M_n$ under linear normalizations. More specifically if for some sequences of normalizing constants $a_n > 0$, $b_n$, $a_n(M_n - b_n)$ has a limiting distribution function $G(x)$, then $G$ must have one of just three possible "forms". This basic, classical result will be proved in Chapter 1. We shall refer to this result as Gnedenko's Theorem, since Gnedenko's paper (1943) gave the first thorough treatment, even though the three extreme value types were previously discovered by Fréchet and Fisher & Tippett.

The second basic result given in Chapter 1 is almost trivial in the independent context, and gives a simple necessary and sufficient conditions under which $P\{M_n \le u_n\}$ converges for a given sequence of constants $\{u_n\}$. This result will play an important role in the dependent case (where it is not as trivial but still true under appropriate conditions).

In Chapter 1 we shall also discuss similar questions for the k - th largest $M_n^{(k)}$, of $\xi_1, \ldots, \xi_n$ where k is fixed, or increasing with n in certain ways. (The case where k increases is described for completeness only and will not be discussed in the later dependent context.)

The task of Chapter 2 is to generalize basic results concerning the maximum $M_n$, to apply to stationary sequences. As will be seen, the generalization follows in a rather complete way under certain natural restrictions limiting the dependence structure of the sequence. In particular it is shown that under these restrictions the limit laws in such dependent cases are precisely the same as the classical ones, and indeed, in a given dependent case, that the same limiting law applies as would if the sequence were *independent*, with the same marginal distribution.

In Chapter 3 this theory is applied to the case of stationary *normal* sequences. It is found there that the dependence conditions required are satisfied under very weak restrictions on the covariances associated with the sequence.

Chapter 4 is concerned with the other topic of Chapter 1 - namely the properties of the k - th largest $M_n^{(k)}$, of the $\xi_i$. The discussion is approached through a consideration of the "point process of exceedances of a level $u_n$" by the sequence $\xi_1, \xi_2, \ldots$ . This provides what we consider to be a helpful and illuminating viewpoint. In particular a simple convergence theorem shows the Poisson nature of the exceedances of high levels, leading to the desired generalizations of the classical results for $M_n^{(k)}$.

As noted, the discussion in Chapters 2 - 4 generalizes the classical case described in Chapter 1, showing that the classical results still apply if the dependence is not "too strong". It is, of course, interesting to investigate the situation under very strong dependence assumptions. There is no complete theory for this case but such questions are explored for normal sequences in Chapter 5, where a variety of different limiting results are found.

# CHAPTER 1

## CLASSICAL EXTREME VALUE THEORY

Let $\xi_1, \xi_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables (r.v.'s), and write $M_n$ for the *maximum* of the first $n$, i.e.

$$(1.1)\qquad M_n = \max(\xi_1, \xi_2, \ldots, \xi_n).$$

Then much of "classical" extreme value theory deals with the distribution of $M_n$, and especially with its properties as $n \to \infty$. All results obtained for maxima of course lead to analogous results for minima through the obvious relation $m_n = \min(\xi_1, \ldots, \xi_n) = -\max(-\xi_1, \ldots, -\xi_n)$. With a few exceptions (e.g. Chapter 11) we shall therefore not discuss minima explicitly in this work.

There is, of course, no difficulty in writing down the distribution function (d.f.) of $M_n$ exactly, in this situation; it is

$$(1.2)\qquad P\{M_n \le x\} = P\{\xi_1 \le x, \xi_2 \le x, \ldots, \xi_n \le x\} = F^n(x),$$

where $F$ denotes the common d.f. of the $\xi_i$. Much of the "Statistics of Extremes" (as is the title of Gumbel's book (1958)) deals with the distribution of $M_n$ in a variety of useful cases, and with a multitude of related questions (for example concerning other order statistics, range of values and so on).

In view of such a satisfactory edifice of theory in finite terms, one may question the desirability of probing for asymptotic results. One reason for such a study appears to us to especially justify it. In simple central limit theory, one obtains an asymptotic normal distribution for the *sum* of many i.i.d. random variables whatever their common original d.f. Indeed one does not have to know the d.f. too precisely to "apply" the asymptotic theory. A similar situation holds in extreme value theory, and, in fact, a non-degenerate asymptotic distribution of $M_n$ (normalized) can take one of just three possible general forms, regardless of the original d.f. $F$. Further, it is not necessary to know the detailed

nature of $F$, to know which limiting form (if any) it gives rise to, i.e. to which "domain of attraction" it belongs. In fact this is determined just by the behaviour of the tail of $F(x)$ for large $x$, and so a good deal may be said about the asymptotic properties of the maximum, based on rather limited knowledge of the properties of $F$.

The central result - here referred to as "Gnedenko's Theorem" - was discorered first by Fisher and Tippet (1928) and later proved in complete generality by Gnedenko (1943). We shall prove this result in the i.i.d. context (Theorem 1.7), using a recent, simple approach due to de Haan (1976), and later extend it to dependent situation (e.g. Theorem 2.4).

We shall be concerned with conditions under which, for suitable normalizing constants $a_n > 0$, $b_n$,

(1.3) $\quad P\{a_n(M_n - b_n) \le x\} \overset{w}{\to} G(x)$

(by which we mean that convergence occurs at continuity points of $G$ – though we shall see later that the $G$'s of interest are all continuous). In particular we shall be interested in determining which d.f.'s $G$ may appear as such a limit. By (1.2), (1.3) may be written as

(1.4) $\quad F^n(a_n^{-1}x + b_n) \overset{w}{\to} G(x)$,

where again the notation $\overset{w}{\to}$ denotes convergence at continuity points of the limiting function.

If (1.4) holds for some sequences $a_n > 0$, $b_n$, we shall say that $F$ belongs to the (i.i.d.) *domain of attraction (for maxima)* of $G$ and write $F \in D(G)$. Equivalently, of course, we may write $a_n$ instead of $a_n^{-1}$ in (1.4) without changing the definition of $D(G)$, and we shall do so without comment where typographically convenient.

A central result to be used in the proof of Gnedenko's Theorem is a general theorem of Khintchine concerning convergence of distribution functions which – for completeness – we shall prove here (Lemma 1.3). It will be convenient to first define inverses of d.f.'s and other monotone func-

tions and to give a few of their simplest properties which we shall use in Lemma 1.3 and subsequently.

DEFINITION If $\psi(x)$ is a non-decreasing right continuous function, we define an inverse function $\psi^{-1}$ (as in de Haan (1976)) by

$$\psi^{-1}(y) = \inf\{x;\ \psi(x) \geq y\}.$$

LEMMA 1.1 (i) For $\psi$ as above, $a > 0$, $b$, $c$ constants, and $H(x) = \psi(ax+b) - c$, then $H^{-1}(y) = a^{-1}(\psi^{-1}(y+c) - b)$.

(ii) For $\psi$ as above, if $\psi^{-1}$ is continuous, then $\psi^{-1}(\psi(x)) = x$.

(iii) If $G$ is a non-degenerate d.f., there exist $y_1 < y_2$ such that $-\infty < G^{-1}(y_1) < G^{-1}(y_2) < \infty$.

PROOF (i) We have

$$\begin{aligned} H^{-1}(y) &= \inf\{x;\ \psi(ax+b) - c \geq y\} \\ &= a^{-1}(\inf\{(ax+b);\ \psi(ax+b) \geq y+c\} - b) \\ &= a^{-1}(\psi^{-1}(y+c) - b), \end{aligned}$$

as required.

(ii) From the definition of $\psi^{-1}$ it is clear that $\psi^{-1}(\psi(x)) \leq x$. If strict inequality holds for some $x$, the definition of $\psi^{-1}$ shows the existence of $z < x$ with $\psi(z) \geq \psi(x)$ and hence $\psi(z) = \psi(x)$ since $\psi$ is non-decreasing. For $y = \psi(z) = \psi(x)$ we have $\psi^{-1}(y) \leq z$ whereas for $y > \psi(z) = \psi(x)$ we have $\psi^{-1}(y) \geq x$, contradicting the continuity of $\psi^{-1}$. Hence $\psi^{-1}(\psi(x)) = x$ as asserted.

(iii) If $G$ is non-degenerate there exist $x_1' < x_2'$ such that $0 < G(x_1') = y_1 < G(x_2') = y_2 \leq 1$. Clearly $x_1 = G^{-1}(y_1)$, $x_2 = G^{-1}(y_2)$ are both finite. Also $G^{-1}(y_2) \geq x_1'$ and equality would require $G(z) \geq y_2$ for all $z > x_1'$ so that $G(x_1') = G(x_1'+0) \geq y_2$, contradicting $G(x_1') = y_1$. Thus $G^{-1}(y_2) > x_1' \geq x_1 = G^{-1}(y_1)$ as required. □

COROLLARY 1.2 If $G$ is a non-degenerate d.f. and $a > 0$, $\alpha > 0$, $b$, $\beta$ constants, such that $G(ax+b) = G(\alpha x+\beta)$ for all $x$, then $a = \alpha$ and $b = \beta$.

PROOF Choose $y_1 < y_2$, $-\infty < x_1 < x_2 < \infty$ by (iii) of the lemma, so that $x_1 = G^{-1}(y_1)$, $x_2 = G^{-1}(y_2)$. Taking inverses of $G(ax+b)$, $G(\alpha x+\beta)$ by (i) of the lemma, we have

$$a^{-1}(G^{-1}(y) - b) = \alpha^{-1}(G^{-1}(y) - \beta)$$

for all $y$. Applying this to $y_1$ and $y_2$ in turn we obtain

$$a^{-1}(x_1 - b) = \alpha^{-1}(x_1 - \beta), \; a^{-1}(x_2 - b) = \alpha^{-1}(x_2 - \beta)$$

from which it follows simply that $a = \alpha$ and $b = \beta$. □

We now obtain the promised general result of Khintchine.

LEMMA 1.3 (Khintchine) Let $\{F_n\}$ be a sequence of d.f.'s, and $G$ a non-degenerate d.f. Let $a_n > 0$, $b_n$ be constants such that

(1.5) $\quad F_n(a_n x + b_n) \xrightarrow{w} G(x)$.

Then for some non-degenerate d.f. $G_*$, constants $\alpha_n > 0$, $\beta_n$,

(1.6) $\quad F_n(\alpha_n x + \beta_n) \xrightarrow{w} G_*(x)$

if and only if

(1.7) $\quad a_n^{-1}\alpha_n \to a$, $a_n^{-1}(\beta_n - b_n) \to b$

for some $a > 0$, $b$ and then

(1.8) $\quad G_*(x) = G(ax+b)$.

PROOF By writing $\alpha_n' = a_n^{-1}\alpha_n$, $\beta_n' = a_n^{-1}(\beta_n - b_n)$, $F_n'(x) = F_n(a_n x + b_n)$, we may rewrite (1.5), (1.6), (1.7) as

(1.5)' $\quad F_n'(x) \xrightarrow{w} G(x)$,

(1.6)' $\quad F_n'(\alpha_n' x + \beta_n') \xrightarrow{w} G_*(x)$,

(1.7)' $\quad \alpha_n' \to a$, $\beta_n' \to b$ for some $a > 0$, $b$.

If (1.5)' and (1.7)' hold, then obviously so does (1.6)' with $G_*(x) = G(ax+b)$. Thus (1.5) and (1.7) imply (1.6) and (1.8).

The proof of the lemma will be complete if we show that (1.5)' and (1.6)' imply (1.7)' for then (1.8) will also hold, as above.

Since $G_*$ is assumed non-degenerate, there are two distinct points $x'$, $x''$ (which may be taken to be continuity points of $G_*$) such that $0 < G_*(x') < 1$, $0 < G_*(x'') < 1$.

The sequence $\{\alpha_n' x' + \beta_n'\}$ must be bounded. For if not, a sequence $\{n_k\}$ could be chosen so that $\alpha_{n_k}' x' + \beta_{n_k}' \to \pm\infty$, which by (1.5)' (since $G$ is a d.f.) would clearly imply that the limit of $F_{n_k}'(\alpha_{n_k}' x' + \beta_{n_k}')$ is zero or one – contradicting (1.6)' for $x = x'$. Hence $\{\alpha_n' x' + \beta_n'\}$ is bounded, and similarly so is $\{\alpha_n' x'' + \beta_n'\}$, which together show that the sequences $\{\alpha_n'\}$, $\{\beta_n'\}$ are each bounded.

Thus there are constants $a$, $b$ and a sequence $\{n_k\}$ of integers such that $\alpha_{n_k}' \to a$, $\beta_{n_k}' \to b$ and it follows as above that

$$(1.9) \qquad F_{n_k}'(\alpha_{n_k}' x + \beta_{n_k}') \xrightarrow{w} G(ax + b),$$

whence since by (1.6)', $G(ax + b) = G_*(x)$, a d.f., we must have $a > 0$. On the other hand if another sequence $\{m_k\}$ of integers gave $\alpha_{m_k}' \to a' > 0$, $\beta_{m_k}' \to b'$ we would have $G(a'x + b') = G_*(x) = G(ax + b)$ and hence $a' = a$, $b' = b$ by Corollary 1.2. Thus $\alpha_n' \to a$, $\beta_n' \to b$ as required to complete the proof. □

We shall say that two d.f.'s $G_1$, $G_2$ are of the *same type* if

$$(1.10) \qquad G_2(x) = G_1(ax + b)$$

for some constants $a > 0$, $b$. Then the above lemma shows that if $\{F_n\}$ is a sequence of d.f.'s with $F_n(a_n x + b_n) \xrightarrow{w} G_1$, $F_n(\alpha_n x + \beta_n) \xrightarrow{w} G_2$, ($a_n > 0$, $\alpha_n > 0$) then $G_1$ and $G_2$ are of the same type, provided they are non-degenerate. As pointed out in de Haan (1976) the d.f.'s may clearly be divided into equivalence classes (which we call *types*) by saying that $G_1$ and $G_2$ are equivalent if $G_2(x) = G_1(ax + b)$ for some $a > 0$, $b$.

If $G_1$ and $G_2$ are d.f.'s of the same type ($G_2(x) = G_1(ax + b)$) and $F \in D(G_1)$, i.e. $F^n(a_n x + b_n) \xrightarrow{w} G_1$ for some $a_n > 0$, $b_n$, then (1.7) is satisfied with $\alpha_n = a_n a$, $\beta_n = b_n + a_n b$, so that $F^n(\alpha_n x + \beta_n) \xrightarrow{w} G_2(x)$ by Lemma 1.3, and hence $F \in D(G_2)$. Thus if $G_1$ and $G_2$ are of the same



We see that for i.i.d. random variables any limiting distribution of the k-th largest, $M^{(k)}$ has the form (1.24)

type, $D(G_1) = D(G_2)$. Similarly we may see from the lemma that if $F$ belongs to both $D(G_1)$ and $D(G_2)$, then $G_1$ and $G_2$ are of the same type. Hence $D(G_1)$ and $D(G_2)$ are identical if $G_1$ and $G_2$ are of the same type, and disjoint otherwise. That is the domain of attraction of a d.f. $G$ depends only on the type of $G$.

The following result gives a characterization for the d.f.'s $G$ for which $D(G)$ is not empty – i.e. for the $G$'s which are possible limiting laws for maxima of i.i.d. sequences. We shall say that a non-degenerate d.f. $G$ is *max-stable* if for each $n = 2, 3, \ldots$ there are constants $a_n > 0$, $b_n$ such that $G^n(a_n x + b_n) = G(x)$. (Again, we can use $a_n^{-1}$ instead of $a_n$ in this definition, and do so as convenient.)

<u>LEMMA 1.4</u> (i) A non-degenerate d.f. $G$ is max-stable if and only if there is a sequence $\{F_n\}$ of d.f.'s, and constants $a_n > 0$, $b_n$ such that

(1.11) $$F_n(a_{nk}^{-1}x + b_{nk}) \xrightarrow{w} G^{1/k}(x) \quad \text{as} \quad n \to \infty$$

for each $k = 1, 2, \cdots$

(ii) In particular if $G$ is non-degenerate, $D(G)$ is non-empty if and only if $G$ is max-stable. Then also $G \in D(G)$.

<u>PROOF</u> (i) If $G$ is non-degenerate, so is $G^{1/k}$ for each $k$, and if (1.11) holds for each $k$, Lemma 1.3 (with $a_n^{-1}$ *for* $a_n$) implies that $G^{1/k}(x) = G(\alpha_k x + \beta_k)$ for some $\alpha_k > 0$, $\beta_k$, so that $G$ is max-stable. Conversely if $G$ is max-stable and $F_n = G^n$, we have $G^n(a_n^{-1}x + b_n) = G(x)$ for some $a_n > 0$, $b_n$ and

$$F_n(a_{nk}^{-1}x + b_{nk}) = [G^{nk}(a_{nk}^{-1}x + b_{nk})]^{1/k} = (G(x))^{1/k}$$

so that (1.11) follows trivially.

(ii) If $G$ is max-stable, $G^n(a_n x + b_n) = G(x)$ for some $a_n > 0$, $b_n$ so (letting $n \to \infty$) we see that $G \in D(G)$. Conversely if $D(G)$ is non-empty, $F \in D(G)$, say, with $F^n(a_n^{-1}x + b_n) \xrightarrow{w} G(x)$. Hence $F^{nk}(a_{nk}^{-1}x + b_{nk}) \xrightarrow{w} G(x)$, or $F^n(a_{nk}^{-1}x + b_{nk}) \xrightarrow{w} G^{1/k}(x)$. Thus (1.11) holds with $F_n = F^n$ and hence by (i) $G$ is max-stable. □

<u>COROLLARY 1.5</u> If $G$ is max-stable, there exist real functions $a(s) > 0$, $b(s)$ defined for $s > 0$ such that

(1.12) $G^s(a(s)x + b(s)) = G(x)$ all real $x$, $s > 0$.

<u>PROOF</u> Since $G$ is max-stable there exists $F \in D(G)$ by Lemma 1.4. Thus we have $a_n > 0$, $b_n$ such that

$$F^n(a_n x + b_n) \xrightarrow{w} G(x) \quad \text{as} \quad n \to \infty$$

and hence (letting [ ] denote integer part)

$$F^{[ns]}(a_{[ns]}x + b_{[ns]}) \xrightarrow{w} G(x),$$

which implies

$$F^{[ns]/s}(a_{[ns]}x + b_{[ns]}) \xrightarrow{w} G^{1/s}(x).$$

But this is easily seen (e.g. by taking logarithms) to give

$$F^n(a_{[ns]}x + b_{[ns]}) \xrightarrow{w} G^{1/s}(x).$$

Since $G^{1/s}$ is non-degenerate, Lemma 1.3 applies with $\alpha_n = a_{[ns]}$, $\beta_n = b_{[ns]}$ to show that $G(a(s)x + b(s)) = G^{1/s}(x)$ for some $a(s) > 0$, $b(s)$, as required. □

We now show that the max-stable d.f.'s consist of precisely three different kinds of types – from which the classical Theorem of Gnedenko will follow simply. We refer to these as Types I, II, and III, even though Types II and III are really families of types, indexed by a parameter $\alpha > 0$.

<u>THEOREM 1.6</u> Every max-stable distribution is one of the following types:

Type I: $G(x) = \exp(-e^{-x})$ $\quad -\infty < x < \infty$,

Type II: $G(x) = 0$ $\quad x \le 0$

$= \exp(-x^{-\alpha})$ (for some $\alpha > 0$) $\quad x > 0$,

Type III: $G(x) = \exp(-(-x)^{\alpha})$ (for some $\alpha > 0$) $\quad x \le 0$

$= 1$ $\quad x > 0$.

<u>PROOF</u> We follow essentially the proof of de Haan (1976). First, it is trivially checked that each of the above types is max-stable.

Conversely, if $G$ is max-stable then (1.12) holds for all $s > 0$, and all $x$. If $0 < G(x) < 1$, (1.12) gives

$$-s \log G(a(s)x + b(s)) = -\log G(x)$$

so that

$$-\log(-\log G(a(s)x + b(s))) - \log s = -\log(-\log G(x)).$$

If $U(x)$ is an inverse function (as defined previously) of $-\log(-\log G(x)) = \psi(x)$, say, then

$$\psi(a(s)x + b(s)) - \log s = \psi(x)$$

so that by Lemma 1.1(i),

$$(U(x + \log s) - b(s))/a(s) = U(x).$$

Subtracting this for $x = 0$ we have

$$(U(x + \log s) - U(\log s))/a(s) = U(x) - U(0)$$

and by writing $y = \log s$, $a_1(y) = a(e^y)$, $\tilde{U}(x) = U(x) - U(0)$

(1.13) $\tilde{U}(x + y) - \tilde{U}(y) = \tilde{U}(x) a_1(y)$,

for all real $x, y$.

Interchanging $x, y$ and subtracting, we obtain

(1.14) $\tilde{U}(x)(1 - a_1(y)) = \tilde{U}(y)(1 - a_1(x))$.

Two cases are possible, (a) and (b) as follows.

(a) $a_1(y) = 1$ for all $y$, when (1.13) gives

$$\tilde{U}(x + y) = \tilde{U}(x) + \tilde{U}(y).$$

The only monotone increasing solution to this is well known to be simply $\tilde{U}(y) = \rho y$ for some $\rho > 0$, so that $U(y) - U(0) = \rho y$, or

$$\psi^{-1}(y) = U(y) = \rho y + \nu, \quad (\nu = U(0)).$$

Since this is continuous, Lemma 1.1(ii) gives

$$x = \psi^{-1}(\psi(x)) = \rho\psi(x) + \nu$$

or $\psi(x) = (x-\nu)/\rho$ so that $G(x) = \exp(-e^{-(x-\nu)/\rho})$ when $0 < G(x) < 1$. It is easy to see from the max-stable property with $n = 2$ that $G$ can have no jump at any finite (upper or lower) end point and hence has the above form for all $x$, thus being of Type I.

(b) $a_1(y) \neq 1$ for some $y$, when (1.14) gives

(1.15) $$\tilde{U}(x) = \frac{\tilde{U}(y)}{1-a_1(y)}(1-a_1(x)) = c_1(1-a_1(x)), \text{ say,}$$

where $c_1 = \tilde{U}(y)/(1-a_1(y)) \neq 0$ (since $\tilde{U}(y) = 0$ would imply $\tilde{U}(x) = 0$ for all $x$ and hence $U(x) = U(0)$, constant).

From (1.13) we thus obtain

$$c_1(1-a_1(x+y)) - c_1(1-a_1(y)) = c_1(1-a_1(x))\,a_1(y),$$

which gives $a_1(x+y) = a_1(x)\,a_1(y)$. But $a_1$ is monotone (from (1.15)) and the only monotone non-constant solutions of this functional equation have the form $a_1(x) = e^{\rho x}$ for $\rho \neq 0$. Hence (1.15) yields

$$\psi^{-1}(x) = U(x) = \nu + c_1(1-e^{\rho x})$$

(where $\nu = U(0)$). Since $-\log(-\log G(x))$ is increasing, so is $U$, so that we must have $c_1 < 0$ if $\rho > 0$ and $c_1 > 0$ if $\rho < 0$. By Lemma 1.1(ii)

$$x = \psi^{-1}(\psi(x)) = \nu + c_1(1-e^{\rho\psi(x)}) = \nu + c_1(1-(-\log G(x))^{-\rho}),$$

giving, where $0 < G(x) < 1$,

$$G(x) = \exp\left\{-\left(1-\frac{x-\nu}{c_1}\right)^{-1/\rho}\right\}.$$

Again from continuity of $G$ at any finite end ponts we see that $G$ is of Type II or Type III with $\alpha = +1/\rho$ or $-1/\rho$ according as $\rho > 0$ ($c_1 < 0$), or $\rho < 0$ ($c_1 > 0$). □

Gnedenko's Theorem now follows immediately for i.i.d. random variables.

<u>THEOREM 1.7</u> (Gnedenko) Let $M_n = \max(\xi_1, \xi_2, \ldots, \xi_n)$ where $\xi_i$ are i.i.d. random variables. If for some constants $a_n > 0$, $b_n$ we have

(1.16) $$P\{a_n(M_n - b_n) \leq x\} \xrightarrow{w} G(x)$$

for some non-degenerate $G$, then $G$ is one of the three extreme value types listed above.

PROOF (1.16) is just (1.4) where $F$ is the d.f. of each $\xi_i$ which by definition gives $F \in D(G)$. Hence by Lemma 1.4(ii), $G$ is max-stable and by Theorem 1.6, $G$ therefore is one of the three listed types. □

Looking ahead, if $\xi_1, \xi_2, \ldots$ are not necessarily independent, but $M_n = \max(\xi_1, \ldots, \xi_n)$ has an asymptotic distribution $G$ in the sense of (1.16), then (1.11) holds with $k = 1$, where $F_n$ is the d.f. of $M_n$. If one can show that if (1.11) holds for $k = 1$, then it holds for all $k$, it will follow that $G$ is max-stable by Lemma 1.4(i), and hence that $G$ is an extreme value type. Thus our approach when we consider dependent cases will be simply to show that under appropriate assumptions, the truth of (1.11) for $k = 1$ implies its truth for all $k$ from which Gnedenko's Theorem will again follow.

Returning now to the i.i.d. case, we note that Gnedenko's Theorem *assumes* that $a_n(M_n - b_n)$ has a non-degenerate limiting d.f. $G$ and then proves that $G$ must have one of the three stated forms. It is easy to construct i.i.d. sequences $\{\xi_n\}$ for which *no* such $G$ exists. For example this is so if $P\{\xi_n < \lambda\} < 1$ and $P\{\xi_n \le \lambda\} = 1$ for some finite $\lambda$. For suppose $P\{a_n(M_n - b_n) \le x\} \to G(x)$, i.e. $P\{M_n \le u_n\} \to G(x)$, where $u_n = x/a_n + b_n$. If $u_n < \lambda$ for infinitely many values of $n$, then $P\{M_n \le u_n\} \le F^n(\lambda - 0)$ for these $n$, and hence $G(x) = 0$ since $F(\lambda - 0) < 1$. Thus if $G(x) > 0$, then $u_n \ge \lambda$ for all sufficiently large $n$, and this means $F(u_n) = 1$, $P\{M_n \le u_n\} = F^n(u_n) = 1$, so that $G(x) = 1$, and $G$ is degenerate.

Other conditions under which no non-degenerate limit exists may be found in Gnedenko (1943), de Haan (1970). For example the conditions given there include the case where $F(x)$ is a Poisson d.f.

On the positive side, various necessary and sufficient conditions are known concerning the domains of attraction. We shall state these without proof. However first we give some very simple and useful *sufficient* conditions which apply when the d.f. $F$ has a density function. These are due to von Mises and elegant, simple proofs are given in de Haan (1976).

Here we reproduce just one of the three proofs as a sample, and refer the reader to the original paper, de Haan (1976), for the details of the others.

THEOREM 1.8 Suppose that the d.f. $F$ of the r.v.'s of the i.i.d. sequence $\{\xi_n\}$ is absolutely continuous with density $f$. Then sufficient conditions for $F$ to belong to each of the three possible domains of attraction are:

Type I: $f$ has a negative derivative $f'$ for all $x$ in some interval $(x_1, x_0)$, $(x_0 \le \infty)$, $f(x) = 0$ for $x \ge x_0$ and

$$\lim_{x \uparrow x_0} f'(x)(1-F(x))/f^2(x) = -1,$$

Type II: $f(x) > 0$ for all $x \ge x_1$ finite, and

$$\lim_{x \to \infty} x\, f(x)/(1-F(x)) = \alpha > 0,$$

Type III: $f(x) > 0$ for all $x$ in some finite interval $(x_1, x_0)$, $f(x) = 0$ for $x > x_0$ and

$$\lim_{x \uparrow x_0} (x_0 - x)\, f(x)/(1-F(x)) = \alpha > 0.$$

PROOF FOR TYPE II case As noted, complete proofs may be found in de Haan (1976), and we give the Type II case here as a sample. Assume then, that $\lim_{x\to\infty} x\, f(x)/(1-F(x)) = \alpha > 0$ where $f(x) > 0$ for $x \ge x_1$. Writing $\alpha(x) = x\, f(x)/(1-F(x))$ it is immediate that for $x_2 \ge x_1$

$$\int_{x_1}^{x_2} (\alpha(t)/t)\, dt = -\log(1-F(x_2)) + \log(1-F(x_1))$$

so that

$$1-F(x_2) = (1-F(x_1)) \exp\left(-\int_{x_1}^{x_2} (\alpha(t)/t)\, dt\right).$$

Clearly there exists $a_n$ such that $1-F(a_n) = n^{-1}$ and (by writing $x_1 = a_n$, $x_2 = a_n x$ or vice versa according as $x \ge 1$ or $x < 1$), we obtain

$$n(1-F(a_n x)) = \exp\left(-\int_{a_n}^{a_n x} (\alpha(t)/t)\, dt\right) = \exp\left(-\int_1^x (\alpha(a_n s)/s)\, ds\right),$$

which, since $a_n \to \infty$, converges to $e^{-\alpha \log x} = x^{-\alpha}$ as $n \to \infty$. Hence for $x > 0$

$$F^n(a_n x) = \left(1 - \frac{n(1 - F(a_n x))}{n}\right)^n \to e^{-x^{-\alpha}}$$

so that the Type II limit follows. (When $x < 0$, $F^n(a_n x) \le F^n(a_n \delta)$ for any $\delta > 0$ and it follows simply that $\lim F^n(a_n x) = 0$). □

If $F$ is normal then $1 - F(x) \sim f(x)/x$ and $f'(x) = -xf(x)$ so that $f'(x)(1 - F(x))/f^2(x) \to -1$ and a Type I limit applies. If $F(x) = 0$ for $x < 0$ and $F(x) = 1 - \exp(-x^\alpha)$ for $x > 0$, $(\alpha > 0)$, $f(x) = \alpha x^{\alpha-1}\exp(-x^\alpha)$, $f'(x) = (\frac{\alpha-1}{x} - \alpha x^{\alpha-1}) f(x)$ so that

$$f'(x)(1 - F(x))/f^2(x) = (\frac{\alpha-1}{x} - \alpha x^{\alpha-1})/(\alpha x^{\alpha-1}) \to -1,$$

again giving a Type I limit. These examples illustrate the simplicity of the criteria. It is similarly easily checked that the Cauchy distribution gives a Type II limit whereas a uniform distribution yields Type III, and each Type I, II, and III limit belongs to its own domain of attraction as previously observed. Other examples are given in de Haan (1976).

As noted above, various necessary and sufficient conditions are also known. These concern the tail behaviour of the d.f. $F$ and do not assume the existence of a density. We state these (as given in Gnedenko (1943)) without proof in order of increasing complexity.

<u>THEOREM 1.9</u> Necessary and sufficient conditions for the d.f. $F$ of the r.v.'s of the i.i.d. sequence $\{\xi_n\}$ to belong to each of the three types are:

Type II: $\lim_{x \to \infty} \frac{1 - F(x)}{1 - F(kx)} = k^\alpha$, $\alpha > 0$, for each $k > 0$,

Type III: there exists $x_0 < \infty$ such that $F(x_0) = 1$, $F(x) < 1$ for all $x < x_0$ and such that

$$\lim_{h \downarrow 0} \frac{1 - F(x_0 - kh)}{1 - F(x_0 - h)} = k^\alpha \quad \text{for each } k > 0,$$

Type I: there exists a continuous function $A(x)$ such that $\lim_{x \uparrow x_0} A(x) = 0$ where $x_0 (\le \infty)$ is such that $F(x_0) = 1$, $F(x) < 1$ for all $x < x_0$, and such that for all $h$,

$$\lim_{x \uparrow x_0} \frac{1 - F\{x(1 + hA(x))\}}{1 - F(x)} = e^{-h}. \qquad \square$$

Note that all these criteria involve the behaviour of $F$ near the "right hand end point" $x_0$ (finite or infinite) where $F$ becomes 1. Note also that if $F(x_0) = 1$ and $F(x) < 1$ for all $x < x_0$ ($x_0$ finite), then it is easily shown that $M_n \to x_0$ with probability one. This does not, in itself, prevent $a_n(M_n - b_n)$ from having a non-degenerate limiting distribution (unless, of course, $F(x_0 - 0) < 1$). The point is that even though $M_n$ itself will usually tend to a degenerate r.v., a suitable normalization can often give a non-degenerate limit (as is familiar in other contexts, such as central limit theory). In fact, if $F(x_0) = 1$ for some finite $x_0$, there may be a limiting d.f. of either Type I or Type III.

The Type II condition may be interpreted as stating that the tail $\psi(x) = 1 - F(x)$ is regularly varying as $x \to \infty$ in the sense that $\psi(x)$ may be expressed as the product of $x^{-\alpha}$ and a function of slow growth. The Type III condition involves similar considerations in the neighbourhood of the finite point $x_0$ at which $F(x)$ becomes unity.

## Convergence of $P\{M_n \leq u_n\}$ and its consequences

Before obtaining a detailed result for normal sequences, we give a general theorem of a form which will be used in later considerations for dependent sequences. It is related to Theorem 1.7 in that the limiting value of $P\{M_n \leq u_n\}$ is investigated as $n \to \infty$. (A generalization of this result applying to other ordered values than maxima appears later in this chapter — Theorem 1.13).

Here we focus on just one sequence $\{u_n\}$ which may, or may not, be

one of a family of the form $x/a_n + b_n$ as $x$ varies.

<u>THEOREM 1.10</u> Let $\{\xi_n\}$ be an i.i.d. sequence. Let $\tau \geq 0$ and suppose $\{u_n\}$ is a sequence of real numbers such that

(1.17) $1 - F(u_n) = \tau/n + o(1/n)$ as $n \to \infty$.

Then

(1.18) $P\{M_n \leq u_n\} \to e^{-\tau}$ as $n \to \infty$.

Conversely if (1.18) holds for some $\tau \geq 0$, so does (1.17).

<u>PROOF</u> From

$$P\{M_n \leq u_n\} = F^n(u_n) = \{1 - (1 - F(u_n))\}^n$$

it is obvious that (1.18) follows from (1.17). Conversely if (1.18) holds then clearly $F(u_n) \to 1$ and by taking logarithms

$$n \log\{1 - (1 - F(u_n))\} \to -\tau$$

from which (1.17) follows. □

One word of caution is perhaps in order here. If $F$ is a *continuous* d.f., we may always choose $u_n$ so that $1 - F(u_n) = \tau/n$ (at least for $\tau > 0$), and then (1.17) trivially holds. If $F$ is not continuous, it is sensible to choose $u_n$ so that $F(u_n - 0) \leq 1 - \tau/n \leq F(u_n)$, e.g. $u_n = \inf\{x;\ F(x) \geq 1 - \tau/n\}$. However, in such a case it does *not* necessarily follow that (1.17) holds. This may be seen without too much difficulty by taking $F$ to increase only by jumps at 1, 2, 3, ... with $F(j) = 1 - \tau 2^{-j}$.

Theorem 1.10 may be used to obtain the asymptotic form of the distibution of $M_n$ when the $\xi_i$ are i.i.d. standard normal random variables.

<u>THEOREM 1.11</u> If $\{\xi_n\}$ is an i.i.d. (standard) normal sequence of r.v.'s then the asymptotic distribution of $M_n = \max(\xi_1, \ldots, \xi_n)$ is of Type I. Specifically

(1.19) $P\{a_n(M_n - b_n) \leq x\} \to \exp(-e^{-x})$,

where

$$a_n = (2\log n)^{1/2},$$

$$b_n = (2\log n)^{1/2} - \frac{1}{2}(2\log n)^{-1/2}(\log\log n + \log 4\pi).$$

PROOF Write $\tau = e^{-x}$ in (1.17). Then we may take $1-\Phi(u_n) = \frac{1}{n}e^{-x}$, where $\Phi$ denotes the standard normal d.f. Since $1-\Phi(u_n) \sim \phi(u_n)/u_n$ ($\phi$ being the standard normal density) we have $\frac{1}{n}e^{-x}u_n/\phi(u_n) \to 1$ and hence $-\log n - x + \log u_n - \log\phi(u_n) \to 0$, or

$$(1.20)\quad -\log n - x + \log u_n + \frac{1}{2}\log 2\pi + u_n^2/2 \to 0.$$

It follows at once that $u_n^2/(2\log n) \to 1$ and hence

$$2\log u_n - \log 2 - \log\log n \to 0$$

or

$$\log u_n = \frac{1}{2}(\log 2 + \log\log n) + o(1).$$

Putting this in (1.20) we obtain

$$\frac{u_n^2}{2} = x + \log n - \frac{1}{2}\log 4\pi - \frac{1}{2}\log\log n + o(1)$$

or

$$u_n^2 = 2\log n\left\{1 + \frac{x - \frac{1}{2}\log 4\pi - \frac{1}{2}\log\log n}{\log n} + o\left(\frac{1}{\log n}\right)\right\}$$

and hence

$$u_n = (2\log n)^{1/2}\left\{1 + \frac{x - \frac{1}{2}\log 4\pi - \frac{1}{2}\log\log n}{2\log n} + o\left(\frac{1}{\log n}\right)\right\},$$

so that

$$u_n = \frac{x}{a_n} + b_n + o((\log n)^{-1/2}) = \frac{x}{a_n} + b_n + o(a_n^{-1}).$$

Hence, since by (1.18) we have $P\{M_n \le u_n\} \to \exp(-e^{-x})$ where $\tau = e^{-x}$,

$$P\{M_n \le x/a_n + b_n + o(a_n^{-1})\} \to \exp(-e^{-x})$$

or

$$P\{a_n(M_n - b_n) + o(1) \le x\} \to \exp(-e^{-x})$$

from which (1.19) follows, as required. □

While there are some computational details in this derivation, there is no difficulty of any kind and (1.19) thus follows in a very simple way from the (itself simple) result (1.18). The same arguments can, and will later, be adapted to a continuous time context.

Poisson nature of exceedances, and implications for k - th maxima

We can look at the choice of $u_n$ which makes (1.17) hold in a slightly different light. Let us regard $u_n$ as a "level" (typically becoming higher with $n$) and say that an *exceedance* of the level $u_n$ by the sequence occurs at "time" $i$ if $\xi_i > u_n$. The probability of such an exceedance is clearly $1 - F(u_n)$ and hence the mean number of exceedances by $\xi_1, \ldots, \xi_n$ is $n(1 - F(u_n)) \to \tau$. That is the choice of $u_n$ is made so that the mean number of exceedances by $\xi_1, \ldots, \xi_n$ is approximately constant. We shall pursue this theme further now in developing Poisson properties of the exceedances. In the following, $S_n$ will denote the number of exceedances of a level $u_n$ by $\xi_1, \ldots, \xi_n$.

THEOREM 1.12 If $\{\xi_n\}$ is an i.i.d. sequence, and if $\{u_n\}$ satisfies (1.17), then $S_n$ is asymptotically Poisson, i.e. for $k = 0, 1, 2, \ldots,$

(1.21) $$P\{S_n \le k\} \to e^{-\tau} \sum_{j=0}^{k} \tau^j/j!$$

Conversely if (1.21) holds for any one fixed $k$, then (1.17) holds (and (1.21) thus holds for all $k$).

PROOF We shall show that if $S_n$ is *any* binomial r.v. with parameters $n$, $p_n$ and $0 \le \tau < \infty$ then (1.21) holds if and only if $np_n \to \tau$. The result will then follow in this particular case with $p_n = 1 - F(u_n)$.

If $S_n$ is binomial and $np_n \to \tau$, then (1.21) follows at once from

the standard Poisson limit for the binomial distribution.

Conversely if (1.21) holds for some $k$ but $np_n \not\to \tau$, there exists $\tau' \neq \tau$, $0 \le \tau' \le \infty$, and a subsequence $\{n_\ell\}$ such that $n_\ell p_{n_\ell} \to \tau'$. If $\tau' < \infty$, the Poisson limit for the binomial distribution shows that $P\{S_{n_\ell} \le k\} \to e^{-\tau'} \Sigma_{j=0}^{k} \tau'^{j}/j!$ as $\ell \to \infty$, which contradicts (1.21) since the function $e^{-x} \Sigma_{j=0}^{k} x^k/k!$ is strictly decreasing in $x \ge 0$ and hence 1 - 1. On the other hand if $\tau' = \infty$,

$$P\{S_n \le k\} = \sum_{r=0}^{k} \binom{n}{r} p_n^r (1-p_n)^{n-r} \le \sum_{r=0}^{k} (np_n)^r (1-p_n)^{n-r}.$$

But $1 - x \le e^{-x}$ for $0 \le x \le 1$ so that

$$(np_n)^r (1-p_n)^{n-r} \le (np_n)^r e^{-(n-r)p_n}$$

which, for each fixed $r$, tends to zero as $n \to \infty$ through the sequence of values $\{n_\ell\}$, since $n_\ell p_{n_\ell} \to \infty$. Thus $P\{S_{n_\ell} \le k\} \to 0$ as $\ell \to \infty$, contradicting the non-zero limit (1.21). Hence we must have $np_n \to \tau$, as asserted. □

We note incidentally, that the random variable $S_n$ used above is asymptotically the same as the number of *upcrossings* of the level $u_n$ between 1 and $n$. Thus we may obtain a Poisson limit for the number of such upcrossings. Such Poisson properties of upcrossings will play an important role when we consider continuous time processes.

## Asymptotic distribution of k - th largest values

If $M_n^{(k)}$ denotes the k - th largest among $\xi_1, \xi_2, \ldots, \xi_n$, then clearly the event $\{M_n^{(k)} \le u_n\}$ is the same as the event $\{S_n < k\}$. By using this we may at once obtain the following restatement of Theorem 1.12.

THEOREM 1.13 Let $\{\xi_n\}$ be an i.i.d. sequence. If $\{u_n\}$ satisfies (1.17) then

(1.22) $P\{M_n^{(k)} \leq u_n\} \to e^{-\tau} \sum_{s=0}^{k-1} \tau^s/s!,$

$k = 1, 2, \ldots$ Conversely if (1.22) holds for some fixed $k$ then (1.17) holds and so does (1.22) for all $k = 1, 2, \ldots$ □

We may further restate this result to give the asymptotic distribution of $M_n^{(k)}$ in terms of that for $M_n (= M_n^{(1)})$.

THEOREM 1.14 Suppose that

(1.23) $P\{a_n(M_n - b_n) \leq x\} \xrightarrow{w} G(x)$

for some non-degenerate (and hence Type I, II, or III) d.f. $G$. Then, for each $k = 1, 2, \ldots,$

(1.24) $P\{a_n(M_n^{(k)} - b_n) \leq x\} \xrightarrow{w} G(x) \sum_{s=0}^{k-1} (-\log G(x))^s/s!$

where $G(x) > 0$ *(and zero where* $G(x) = 0$).

Conversely if for some fixed $k$,

(1.25) $P\{a_n(M_n^{(k)} - b_n) \leq x\} \xrightarrow{w} H(x)$

for some non-degenerate $H$, then $H(x)$ must be of the form on the right hand side of (1.24), where (1.23) holds with the same $G$, $a_n$, $b_n$. (Hence (1.24) holds for all $k$.)

PROOF If (1.23) holds and $G(x) > 0$ then (1.22) holds with $k = 1$, $\tau = -\log G(x)$, so that by Theorem 1.13 (1.22) holds for all $k$, i.e. (1.24) follows. The case $G(x) = 0$ follows, using continuity.

Conversely if (1.25) holds, for some fixed $k$, and $x$ is such that $H(x) > 0$, we may clearly find $\tau \geq 0$ such that

$$H(x) = e^{-\tau} \sum_{s=0}^{k-1} \tau^s/s!,$$

since this function decreases continuously from 1 to 0 as $\tau$ increases. Thus (1.22) holds for this $k$ and hence, by Theorem 1.13 for all $k$ including $k = 1$, which gives (1.23) with $\tau = -\log G(x)$ (non-degeneracy of $G$ being clear). □

We see that for i.i.d. random variables any limiting distribution of the k - th largest, $M_n^{(k)}$, has the form (1.24) based on the same d.f. G as applied to the maximum, and moreover that the normalizing constants are the same for all k including $k = 1$ (the maximum itself). Thus we have a complete description of the possible non-degenerate limiting laws.

## Joint asymptotic distribution of the largest maxima

The asymptotic distribution of the k - th largest maximum was obtained above by considering the number of exceedances of a level $u_n$ by $\xi_1, \ldots, \xi_n$. Similar arguments can, and will presently, be adapted to prove convergence of the joint distribution of several large maxima.

Let the levels $u_n^{(1)} \geq \ldots \geq u_n^{(r)}$ satisfy

$$(1.26)\qquad \begin{aligned} 1 - F(u_n^{(1)}) &= \tau_1/n + o(1/n), \\ &\vdots \\ 1 - F(u_n^{(r)}) &= \tau_r/n + o(1/n), \end{aligned}$$

and define $S_n^{(k)}$ to be the number of exceedances of $u_n^{(k)}$ by $\xi_1, \ldots, \xi_n$.

<u>THEOREM 1.15</u> Suppose that $\{\xi_n\}$ is an i.i.d. sequence and that $\{u_n^{(k)}\}$, $k = 1, \ldots, r$, satisfy (1.26). Then, for $k_1 \geq 0, \ldots, k_r \geq 0$,

$$(1.27)\qquad P\{S_n^{(1)} = k_1,\ S_n^{(2)} = k_1 + k_2,\ \ldots,\ S_n^{(r)} = k_1 + \ldots + k_r\}$$

$$\to \frac{\tau_1^{k_1}}{k_1!} \cdot \frac{(\tau_2 - \tau_1)^{k_2}}{k_2!} \cdot \ldots \cdot \frac{(\tau_r - \tau_{r-1})^{k_r}}{k_r!}\, e^{-\tau_r}$$

as $n \to \infty$.

<u>PROOF</u> Writing $p_{n,k} = 1 - F(u_n^{(k)})$ for the probability that $\xi_1$ exceeds $u_n^{(k)}$, it is easy to see that the left-hand side of (1.27) equals

$$(1.28)\qquad \binom{n}{k_1} p_{n,1}^{k_1} \binom{n-k_1}{k_2} (p_{n,2} - p_{n,1})^{k_2} \cdot \ldots \cdot \binom{n-k_1-\ldots-k_{r-1}}{k_r} (p_{n,r} - p_{n,r-1})^{k_r} \cdot (1 - p_{n,r})^{n-k_1-\ldots-k_r}.$$

From (1.26) it follows in turn that

$$\binom{n}{k_1} p_{n,1}^{k_1} = n\cdot\ldots\cdot(n-k_1+1)\, p_{n,1}^{k_1}/k_1! \to \tau_1^{k_1}/k_1!,$$

$$\binom{n-k_1-\ldots-k_{\ell-1}}{k_\ell}\left(p_{n,\ell}-p_{n,\ell-1}\right)^{k_\ell}$$

$$= (n-k_1-\ldots-k_{\ell-1})\cdot\ldots\cdot(n-k_1-\ldots-k_\ell+1)(p_{n,\ell}-p_{n,\ell-1})^{k_\ell}/k_\ell!$$

$$\to (\tau_\ell-\tau_{\ell-1})^{k_\ell}/k_\ell!$$

for $2 \le \ell \le r$, and that

$$(1-p_{n,r})^{n-k_1-\ldots-k_r} \to e^{-\tau_r},$$

and thus (1.27) is an immediate consequence of (1.26) and (1.28). □

Clearly

$$(1.29)\quad P\{M_n^{(1)} \le u_n^{(1)}, \ldots, M_n^{(r)} \le u_n^{(r)}\} = P\{S_n^{(1)} = 0, S_n^{(2)} \le 1, \ldots, S_n^{(r)} \le r-1\},$$

and thus the joint asymptotic distribution of the $r$ largest maxima can be obtained directly from Theorem 1.15. In particular, if the distribution of $a_n(M_n^{(1)}-b_n)$ converges, then it follows not only that $a_n(M_n^{(k)}-b_n)$ converges in distribution for $k = 2, 3, \ldots$ as was seen above, but also that the joint distribution of $a_n(M_n^{(1)}-b_n), \ldots, a_n(M_n^{(r)}-b_n)$ converges. This is of course completely straightforward, but since the form of the limiting distribution becomes somewhat complicated if more than two maxima are considered, we state the result only for the two largest maxima.

<u>THEOREM 1.16</u> Suppose that

$$(1.30)\quad P\{a_n(M_n^{(1)}-b_n) \le x\} \xrightarrow{w} G(x)$$

for some non-degenerate (and hence Type I, II, or III) d.f. $G$. Then, for $x_1 > x_2$,

(1.31) $P\{a_n(M_n^{(1)} - b_n) \leq x_1,\ a_n(M_n^{(2)} - b_n) \leq x_2\}$

$$\xrightarrow{w} G(x_2)\{\log G(x_1) - \log G(x_2) + 1\},$$

when $G(x_2) > 0$ (and to zero when $G(x_2) = 0$).

PROOF We have to prove that

$$P\{M_n^{(1)} \leq u_n^{(1)},\ M_n^{(2)} \leq u_n^{(2)}\}$$

converges if $u_n^{(1)} = x_1/a_n + b_n$ and $u_n^{(2)} = x_2/a_n + b_n$. If (1.30) holds and $G(x_2) > 0$ then $1 - F(u_n^{(1)}) \sim \tau_1/n$ and $1 - F(u_n^{(2)}) \sim \tau_2/n$, where $\tau_1 = -\log G(x_1)$, $\tau_2 = -\log G(x_2)$. Hence, by Theorem 1.15

$$\begin{aligned} P\{S_n^{(1)} = 0,\ S_n^{(2)} \leq 1\} &= P\{S_n^{(1)} = 0,\ S_n^{(2)} = 0\} + P\{S_n^{(1)} = 0,\ S_n^{(2)} = 1\} \\ &\to e^{-\tau_2} + (\tau_2 - \tau_1)e^{-\tau_2} = e^{-\tau_2}(\tau_2 - \tau_1 + 1), \end{aligned}$$

which by (1.29) proves (1.31). □

Increasing ranks

The results above apply to the $k$-th largest $M_n^{(k)}$ of $\xi_1, \xi_2, \ldots, \xi_n$, when $k$ is fixed. We refer to this as the case of *fixed ranks* (or *extreme* order statistics). It is also of interest to consider cases where $k = k_n \to \infty$ as $n \to \infty$ and we shall refer to this as the case of *increasing* ranks. Two particular rates of convergence are of special interest:

(i) $k_n/n \to \theta$ $(0 < \theta < 1)$, which we shall call the case of *central* ranks,

(ii) $k_n \to \infty$ but $k_n/n \to 0$, which will be called the *intermediate* rank case.

For the consideration of fixed ranks it was useful to define levels $\{u_n\}$ satisfying (1.17), i.e. $n(1 - F(u_n)) \to \tau$. In the case where $k_n \to \infty$ we shall find that the appropriate restrictions are that $nF(u_n)(1 - F(u_n)) \to \infty$ and, writing $p_n = 1 - F(u_n)$,

$$(1.32)\qquad \frac{k_n - n\,p_n}{(n\,p_n(1-p_n))^{1/2}} \to \tau$$

for a fixed constant $\tau$, or equivalently (as we shall see),

$$(1.33)\qquad n - k_n \to \infty,\quad \frac{k_n - n\,p_n}{(k_n(1-k_n/n))^{1/2}} \to \tau.$$

Theorem 1.11 now has the following counterpart (in which $\Phi$ will denote the standard normal d.f., and $S_n$ is again the numbers of exceedances of a level $u_n$ by $\xi_1, \xi_2, \ldots, \xi_n$).

<u>THEOREM 1.17</u> With the above notation, let $k_n \to \infty$, write $p_n = 1 - F(u_n)$, and suppose $n\,p_n(1-p_n) \to \infty$. If $\{u_n\}$ satisfies (1.32), then

$$(1.34)\qquad P\{S_n \le k_n\} \to \Phi(\tau) \quad \text{as} \quad n \to \infty.$$

*Conversely if (1.34) holds so does (1.32).*

In the above statements (1.32) can be replaced by the equivalent condition (1.33).

<u>PROOF</u> We may write $S_n = \sum_1^n \chi_i$ where $\chi_i$ = 1 or 0 according as $\xi_i > u_n$ or $\xi_i \le u_n$. The $\chi_i$ are thus i.i.d. with $P\{\chi_i = 1\} = p_n = 1 - P\{\chi_i = 0\}$. It follows from the Berry-Esseen bound that

$$\left|P\{S_n \le k_n\} - \Phi\left(\frac{k_n - n\,p_n}{(n\,p_n(1-p_n))^{1/2}}\right)\right| \le C/(n\,p_n(1-p_n))^{1/2},$$

which tends to zero since $n\,p_n(1-p_n) \to \infty$. The main result follows since

$$\Phi\left(\frac{k_n - n\,p_n}{(n\,p_n(1-p_n))^{1/2}}\right) \to \Phi(\tau)$$

if and only if $(k_n - n\,p_n)/(n\,p_n(1-p_n))^{1/2} \to \tau$, ($\Phi$ and its inverse function both being continuous).

Finally, that (1.32) implies (1.33) follows by writing $k_n = n\,p_n + \tau(n\,p_n(1-p_n))^{1/2}(1+o(1))$, and noting that this implies $k_n \sim n\,p_n$ and $(n-k_n) \sim n(1-p_n)$. Similarly (1.33) implies (1.32). □

Corresponding to Theorems 1.13, 1.14, we thus have the following results.

__THEOREM 1.18__ With the above notation, suppose that $k_n \to \infty$, $n p_n(1 - p_n) \to \infty$ $(p_n = 1 - F(u_n))$. If (1.32) or (1.33) holds, then

(1.35) $P\{M_n^{(k_n)} \leq u_n\} \to \Phi(\tau)$.

Conversely if (1.35) holds so do (1.32) and (1.33) □

__THEOREM 1.19__ Again with the above notation, suppose that (1.32) or (1.33) holds with $u_n = u_n(x) = x/a_n + b_n$ $(\tau = \tau(x))$ for some sequences $\{a_n > 0\}$, $\{b_n\}$. Then

(1.36) $P\{a_n(M_n^{(k_n)} - b_n) \leq x\} \xrightarrow{w} H(x)$

where $H(x) = \Phi(\tau(x))$. Conversely if (1.36) holds for some non-degenerate d.f. $H$, then we have $H(x) = \Phi(\tau(x))$ where (1.32) and (1.33) hold with $u_n = x/a_n + b_n$, $\tau = \tau(x)$. □

## Central ranks

The case of central ranks, where $k_n/n \to \theta$ $(0 < \theta < 1)$ has been studied in Smirnov (1952). While we shall have little to say about this in later chapters, for the sake of completeness a few basic facts for the i.i.d. sequence will be discussed here. First we note that it is possible for two sequences $\{k_n\}$, $\{k_n'\}$ with $\lim k_n/n = \lim k_n'/n$ to lead to *different* non-degenerate limiting d.f.'s for $M_n^{(k_n)}$, $M_n^{(k_n')}$. Specifically, as shown in Smirnov (1952), we may have $k_n/n \to \theta$, $k_n'/n \to \theta$ and

(1.37) $P\{a_n(M_n^{(k_n)} - b_n) \leq x\} \xrightarrow{w} H(x)$,

(1.38) $P\{a_n'(M_n^{(k_n')} - b_n') \leq x\} \xrightarrow{w} H'(x)$,

where $a_n > 0$, $b_n$, $a_n' > 0$, $b_n'$ are constants and $H(x)$, $H'(x)$ are non-degenerate d.f.'s of different "type". However this is not possible if

(1.39) $\sqrt{n}\left(\frac{k_n}{n} - \theta\right) \to 0$

as the following result shows.

LEMMA 1.20 Suppose that (1.37) and (1.38) hold where $H$, $H'$ are non-degenerate and $k_n$, $k_n'$ both satisfy (1.39). Then $H$ and $H'$ are of the same "type", i.e. $H'(x) = H(ax + b)$ for some $a > 0$, $b$.

PROOF If the terms of the i.i.d. sequence $\xi_1, \xi_2, \ldots,$ have d.f. $F$, then by Theorem 1.19

$$(1.40) \qquad \frac{k_n - n(1 - F(x/a_n + b_n))}{(k_n(1 - k_n/n))^{1/2}} \to \tau(x)$$

where $H(x) = \Phi(\tau(x))$. By (1.39) we then have

$$\sqrt{n}\,\frac{\theta - (1 - F(x/a_n + b_n))}{(\theta(1-\theta))^{1/2}} \to \tau(x).$$

Again by (1.39), with $k_n$ *replaced by* $k_n'$, we must therefore have that (1.40) holds with $k_n'$ replacing $k_n$, and hence by Theorem 1.19, that

$$P\{a_n(M_n^{(k_n')} - b_n) \le x\} \to \Phi(\tau(x)) = H(x).$$

But if $H_n$ is the d.f. of $M_n^{(k_n')}$ this says that $H_n(x/a_n + b_n) \xrightarrow{w} H(x)$, *whereas also* $H_n(x/a_n' + b_n') \xrightarrow{w} H'(x)$ by (1.38). Thus, by Lemma 1.3, $H$ and $H'$ are of the same type as required. □

It turns out that for sequences $\{k_n\}$ satisfying (1.39) just four types of limiting distributions $H$ satisfying (1.37) are possible for $M_n^{(k_n)}$. For completeness we state this here as a theorem – and refer to Smirnov (1952) for proof.

THEOREM 1.21 If the central rank sequence $\{k_n\}$ satisfies (1.39) the only possible non-degenerate d.f.'s $H$ for which (1.37) holds are

1. $$H(x) = 0, \qquad x < 0$$
$$= \Phi(cx^{\alpha}) \qquad x \ge 0 \quad (c > 0,\ \alpha > 0),$$

2. $$H(x) = \Phi(-c|x|^{\alpha}) \qquad x < 0 \quad (c > 0,\ \alpha > 0)$$
$$= 1 \qquad x \ge 0,$$

3. $$H(x) = \Phi(-c_1|x|^{\alpha}) \quad x < 0$$
$$= \Phi(c_2 x^{\alpha}) \quad x \geq 0 \quad (c_1 > 0,\ c_2 > 0,\ \alpha > 0),$$

4. $$H(x) = 0 \quad x < -1$$
$$= 1/2 \quad -1 \leq x < 1$$
$$= 1 \quad x \geq 1. \qquad \square$$

Intermediate ranks

By an *intermediate rank* sequence we mean a sequence $\{k_n\}$ such that $k_n \to \infty$ but $k_n = o(n)$. The general theory for increasing ranks applies, with some slight simplification. For example we may rephrase (1.33) as

$$p_n = \frac{k_n}{n} - \tau\frac{k_n^{1/2}}{n} + o\left(\frac{k_n^{1/2}}{n}\right).$$

The following result of Wu (1966) gives the possible normalized limiting d.f.'s of $M_n^{(k_n)}$ when $k_n$ is non-decreasing.

THEOREM 1.22 If $\xi_1, \xi_2, \ldots$ are i.i.d. and $\{k_n\}$ is a non-decreasing intermediate rank sequence, and if there are constants $a_n > 0$, $b_n$ such that

$$P\{a_n(M_n^{(k_n)} - b_n) \leq x\} \xrightarrow{w} H(x)$$

for a non-degenerate d.f. $H$, then $H$ has one of the three forms

$$H_1(x) = \Phi(-a\log|x|) \quad x < 0 \quad (a > 0)$$
$$= 1 \quad x \geq 0,$$

$$H_2(x) = 0 \quad x \leq 0 \quad (a > 0)$$
$$= \Phi(a\log x) \quad x > 0,$$

$$H_3(x) = \Phi(x) \quad -\infty < x < \infty. \qquad \square$$

This theorem is rather satisfying, though it does not specify the domains of attraction of the three limiting forms. Some results in this direction have been obtained in Chibisov (1964), Smirnov (1967), and Wu

(1966). However these are highly dependent on the rank sequence $\{k_n\}$. For example a class of rank sequences $\{k_n\}$ such that $k_n \sim \ell^2 n^\theta$ $(0 < \theta < 1)$ are studied in Chibisov (1964). If $F$ is any d.f. it is known that there is at most one pair of $(\ell, \theta)$ such that $F$ belongs to the domain of attraction of $H_1$ and the same statement holds for $H_2$. In addition, there are rank sequences such that only the normal law $H_3$ is a possible limit and, moreover, there are distributions attracted to it for every intermediate rank sequence $\{k_n\}$.

The above brief survey of selected parts of classical extreme value theory is by no means comprehensive or complete. However the topics have been chosen to illustrate the theory, and to provide a starting point for the discussion of *dependent* cases. In the next chapters we shall take up some of these topics directly, but will also consider a number of areas which are related to extreme value theory and have importance in their own right.

## CHAPTER 2

## MAXIMA OF STATIONARY SEQUENCES

There are various ways in which the notion of an i.i.d. sequence may be generalized by "introducing dependence", or allowing the $\xi_n$ to have different distributions, or both. For example an obvious generalization is to consider a sequence which is Markov of some order. Though the consideration of Markov sequences can lead to fruitful results for extremes, it is not the direction we shall pursue here.

We shall keep the assumption that the $\xi_n$ have a common distribution; in fact it will be natural to consider *stationary* sequences, i.e. sequences such that the distributions of $(\xi_{j_1}, \ldots, \xi_{j_n})$ and $(\xi_{j_1+m}, \ldots, \xi_{j_n+m})$ are identical for any choice of $n, j_1, \ldots, j_n$, and $m$. Then we shall assume that the dependence between $\xi_i$ and $\xi_j$ falls off in some specified way as $|i-j|$ increases. This is different from the Markov property where in essence the past, $\{\xi_i; i < n\}$, and the future, $\{\xi_j; j > n\}$, are independent, given the present, $\xi_n$.

The simplest example of the type of restriction we consider, is that of m-dependence, which requires that $\xi_i$ and $\xi_j$ be actually *independent* if $|i-j| > m$.

A more commonly used dependence restriction of this type for stationary sequences is that of *strong mixing* (introduced first in Rosenblatt (1956)). Specifically, the sequence $\{\xi_n\}$ is said to satisfy the strong mixing assumption if there is a function $g(k)$, the "mixing function", tending to zero as $k \to \infty$, and such that

$$|P(A \cap B) - P(A)P(B)| < g(k)$$

when $A \in \mathcal{F}(\xi_1, \ldots, \xi_p)$ and $B \in \mathcal{F}(\xi_{p+k+1}, \xi_{p+k+2}, \ldots)$ for any $p$ and $k$, $\mathcal{F}(\cdot)$ denotes the $\sigma$-field generated by the indicated random variables. Thus when a sequence is mixing, any event $A$ "based on the past up to time $p$" is "nearly independent" of any event $B$ "based on the future from time $p+k+1$ onwards", when $k$ is large. Note that

this mixing condition is uniform in the sense that $g(k)$ does not depend on the particular A and B involved.

The *correlation* between $\xi_i$ and $\xi_j$ is also a (partial) measure of their dependence. Hence another dependence restriction of the same type is $|\mathrm{Corr}(\xi_i, \xi_j)| \leq g(|i-j|)$ where $g(k) \to 0$ as $k \to \infty$. Obviously such a restriction will be most useful if the $\xi_n$ form a normal sequence.

Various results from extreme value theory have been extended to apply under each of the more general restrictions mentioned above. For example Watson (1954) generalized (1.18) to apply to m-dependent stationary sequences. Loynes (1965) considered quite a number of results (including (1.18) and Gnedenko's Theorem) under the strong mixing assumption, for stationary sequences. Berman (1964) used some simple correlation restrictions to obtain (1.19) for stationary normal sequences.

It is obvious that the results of Loynes (1965) and Berman (1964) are related - similar methods being useful in each - but the precise connections are not immediately apparent, due to the different dependence restrictions used. Berman's correlation restrictions are very weak assumptions, leading to sharp results for *normal* sequences. The mixing condition used by Loynes, while being useful in some contexts, is obviously rather restrictive. In this chapter we shall propose a much weaker condition of "mixing type", which first appeared in Leadbetter (1974), and under which, for example, the results of Loynes (1965) will still be true. Further, this condition will be satisfied for stationary normal sequences under Berman's correlation conditions (as we shall see in the next chapter). Hence the relationships between the various results are clarified.

## Mixing conditions for Gnedenko's theorem

In weakening the mixing condition, one notes that the events of interest in extreme value theory are typically those of the form $\{\xi_i \leq u\}$ or their intersections. For example the event $\{M_n \leq u\}$ is just

$\{\xi_1 \leq u, \xi_2 \leq u, \ldots, \xi_n \leq u\}$. Hence one may be led to propose a condition like mixing but only required to hold for events of this type. For example one such natural condition would be the following, which we shall call Condition D. For brevity we will write $F_{i_1 \ldots i_n}(u)$ for $F_{i_1 \ldots i_n}(u, \ldots, u)$, if $F_{i_1 \ldots i_n}(x_1, \ldots, x_n)$ denotes the joint d.f. of $\xi_{i_1}, \ldots, \xi_{i_n}$.

<u>CONDITION D</u> will be said to hold if for any integers $i_1 < \ldots < i_p$ and $j_1 < \ldots < j_q$ for which $j_1 - i_p \geq \ell$, and any real $u$,

$$(2.1) \qquad |F_{i_1 \ldots i_p, j_1 \ldots j_q}(u) - F_{i_1 \ldots i_p}(u) F_{j_1 \ldots j_q}(u)| \leq g(\ell)$$

where $g(\ell) \to 0$ as $\ell \to \infty$.

We shall see that Gnedenko's Theorem - and a number of other results - hold under D. However, while D is a significant reduction of the requirements imposed by mixing, we can do better yet. We shall consider a condition, to be called $D(u_n)$, which will involve a requirement like (2.1) but applying only to a certain sequence of values $\{u_n\}$ and not necessarily to all u-values. More precisely, if $\{u_n\}$ is a given real sequence, we define the condition $D(u_n)$ as follows.

<u>CONDITION $D(u_n)$</u> will be said to hold if for any integers $1 \leq i_1 < \ldots < i_p < j_1 \ldots < j_q \leq n$ for which $j_1 - i_p \geq \ell$, we have

$$(2.2) \qquad |F_{i_1 \ldots i_p, j_1 \ldots j_q}(u_n) - F_{i_1 \ldots i_p}(u_n) F_{j_1 \ldots j_q}(u_n)| \leq \alpha_{n,\ell}$$

where $\alpha_{n,\ell_n} \to 0$ as $n \to \infty$ for some sequence $\ell_n = o(n)$.

Note that for each $n, \ell$ we may replace $\alpha_{n,\ell}$ by the maximum of the left hand side of (2.2) over all allowed choices of i's and j's, to obtain a possibly smaller $\alpha_{n,\ell}$ which is non-increasing in $\ell$ for each fixed $n$, and still satisfies $\alpha_{n,\ell_n} \to 0$ as $n \to \infty$. Note also that for such $\alpha_{n,\ell}$, taken non-increasing in $\ell$ for each fixed $n$, the condition $\alpha_{n,\ell_n} \to 0$ as $n \to \infty$, $\ell_n = o(n)$, may be rewritten as

(2.3) $\alpha_{n,n\lambda} \to 0$ for each $\lambda > 0$.

It is trivially seen that if $\alpha_{n,\ell_n} \to 0$ for some $\ell_n = o(n)$, then (2.3) holds. The converse may be shown by noting that (2.3) implies the existence of an increasing sequence of constants $m_k$ such that $\alpha_{n,n/k} < k^{-1}$ for $n \geq m_k$. If $k_n$ is defined by $k_n = r$ for $m_r \leq n < m_{r+1}$, $r \geq 1$, then $m_{k_n} \leq n$ so that $\alpha_{n,n/k_n} \leq k_n^{-1} \to 0$, and the sequence $\{\ell_n\}$ may be taken to be $\{n/k_n\}$.

Strong mixing implies D, which in turn implies $D(u_n)$ for any sequence $\{u_n\}$. Also $D(u_n)$ is satisfied, for appropriately chosen $\{u_n\}$, by stationary normal sequences under weak conditions, whereas mixing need not be (and it seems likely that D need not be either, though we have not investigated this).

First we give a lemma which demonstrates how Condition $D(u_n)$ gives the "degree of independence" appropriate for our purposes. If E is any set of integers, M(E) will denote $\max\{\xi_j;\ j \in E\}$ (and of course $M(E) = M_n$ if $E = \{1, \ldots, n\}$). It will be convenient to let an "interval" mean any finite set E of consecutive integers $\{j_1, \ldots, j_2\}$, say; its *length* will be taken to be $j_2 - j_1 + 1$. If $F = \{k_1, \ldots, k_2\}$ is another interval with $k_1 > j_2$, we shall say that E and F are *separated by* $k_1 - j_2$.

Throughout, $\{\xi_n\}$ will be a stationary sequence.

<u>LEMMA 2.1</u> Suppose $D(u_n)$ holds for some sequence $\{u_n\}$. Let n, r, k be fixed integers and $E_1, \ldots, E_r$ subintervals of $\{1, \ldots, n\}$ such that $E_i$ and $E_j$ are separated by at least k when $i \neq j$. Then

$$\left| P\left( \bigcap_{j=1}^{r} \{M(E_j) \leq u_n\} \right) - \prod_{j=1}^{r} P\{M(E_j) \leq u_n\} \right| \leq (r-1)\alpha_{n,k}.$$

<u>PROOF</u> This is easily shown inductively. For brevity write $A_j = \{M(E_j) \leq u_n\}$. Let $E_j = \{k_j, \ldots, \ell_j\}$ where (by renumbering if necessary) $k_1 \leq \ell_1 < k_2 \leq \ldots$ . Then

$$|P(A_1 \cap A_2) - P(A_1)P(A_2)|$$
$$= |F_{k_1 \ldots \ell_1, k_2 \ldots \ell_2}(u_n) - F_{k_1 \ldots \ell_1}(u_n)F_{k_2 \ldots \ell_2}(u_n)|$$
$$\leq \alpha_{n,k}$$

since $k_2 - \ell_1 \geq k$. Similarly

$$|P(A_1 \cap A_2 \cap A_3) - P(A_1)P(A_2)P(A_3)|$$
$$\leq |P(A_1 \cap A_2 \cap A_3) - P(A_1 \cap A_2)P(A_3)| + |P(A_1 \cap A_2) - P(A_1)P(A_2)|P(A_3)$$
$$\leq 2\,\alpha_{n,k}$$

since $E_1 \cup E_2 \subseteq \{k_1, \ldots, \ell_2\}$ and $k_3 - \ell_2 \geq k$. Proceeding in this way we obtain the result. □

This lemma shows a degree of independence for maxima on separated intervals. To obtain Gnedenko's Theorem from Lemma 1.4 we need to show that if (1.11) holds for $k = 1$, it holds for all $k$, i.e. that if $P\{a_n(M_n - b_n) \leq x\} \to G(x)$, then $P\{a_{nk}(M_n - b_{nk}) \leq x\} \to G^{1/k}(x)$ for each $k = 2, 3, \ldots$ This will be so if, for each $k = 2, 3, \ldots$

(2.4) $\quad P\{M_{nk} \leq x/a_{nk} + b_{nk}\} - P^k\{M_n \leq x/a_{nk} + b_{nk}\} \to 0$

as $n \to \infty$. Hence we wish to prove (2.4) to obtain Gnedenko's Theorem.

The method used is to divide the interval $\{1, \ldots, n\}$ into $k$ intervals of length $[n/k]$, and use "approximate independence" of the maxima on each via Lemma 2.1 to give a result which implies (2.4). To apply Lemma 2.1, we must shorten each of these intervals to separate them. This leads to the following construction, used for example in Loynes (1965), and given here in a slightly more general form for later use also.

Specifically we suppose that $k$ is a fixed integer, and for any positive integer $n$, write $n' = [n/k]$, (*the integer part of* $n/k$). Thus we have $n'k \leq n < (n'+1)k$. Divide the first $n'k$ integers into $2k$ consecutive intervals, as follows. For large $n$, let $m$ be an integer,

$k < m < n'$, and write $I_1 = \{1, 2, \ldots, n'-m\}$, $I_1^* = \{n'-m+1, \ldots, n'\}$, $I_2, I_2^*, \ldots, I_k, I_k^*$ being defined similarly, alternately having length $n'-m$ and $m$. Finally write

$$I_{k+1} = \{(k-1)n'+m+1, \ldots, kn'\},\ I_{k+1}^* = \{kn'+1, \ldots, kn'+m\}.$$

(Note that $I_{k+1}, I_{k+1}^*$ are defined differently from $I_j, I_j^*$ for $j \le k$.)

The main steps of the approximation are contained in the following lemma. These are, broadly, to show first that the "small" intervals $I_j^*$ can be essentially disregarded, and then to apply Lemma 2.1 to the (now separate) intervals $I_1, \ldots, I_k$. In the following, $\{u_n\}$ is any given sequence (not necessarily of the form $x/a_n + b_n$).

LEMMA 2.2 With the above notation, and assuming $D(u_n)$ holds,

(i) $$0 \le P\left(\bigcap_{j=1}^{k} \{M(I_j) \le u_n\}\right) - P\{M_n \le u_n\} \le (k+1)P\{M(I_1) \le u_n < M(I_1^*)\}$$

(ii) $$\left|P\left(\bigcap_{j=1}^{k} \{M(I_j) \le u_n\}\right) - P^k\{M(I_1) \le u_n\}\right| \le k\,\alpha_{n,m}$$

(iii) $$|P^k\{M(I_1) \le u_n\} - P^k\{M_{n'} \le u_n\}| \le k\,P\{M(I_1) \le u_n < M(I_1^*)\}.$$

Hence, by combining (i), (ii), (iii),

(2.5) $$|P\{M_n \le u_n\} - P^k\{M_{n'} \le u_n\}| \le (2k+1)P\{M(I_1) \le u_n < M(I_1^*)\} + k\,\alpha_{n,m}.$$

PROOF The result (i) follows at once since $\bigcap_{j=1}^{k} \{M(I_j) \le u_n\} \supset \{M_n \le u_n\}$, and their difference implies $M(I_j) \le u_n < M(I_j^*)$ for some $j \le k$ or otherwise $\xi_j \le u_n$ for $1 \le j \le k\,n'$ but $\xi_j > u_n$ for some $j = k\,n'+1, \ldots, k(n'+1)$, which in turn implies $M(I_{k+1}) \le u_n < M(I_{k+1}^*)$, (since $m > k$ and hence $k(n'+1) < k\,n'+m$). Since the probabilities of the events $M(I_j) \le u_n < M(I_j^*)$ are independent of $j$ by stationarity, (i) follows.

The inequality (ii) follows from Lemma 2.1 with $I_j$ for $E_j$, noting that $P\{M(I_j) \le u_n\}$ is independent of $j$.

To obtain (iii) we note that

$$0 \le P\{M(I_1) \le u_n\} - P\{M_{n'} \le u_n\} = P\{M(I_1) \le u_n < M(I_1^*)\}.$$

The result then follows, writing $y = P\{M(I_1) \le u_n\}$ and $x = P\{M_{n'} \le u_n\}$, from the obvious inequalities

$$0 \le y^k - x^k \le k(y-x) \quad \text{when} \quad 0 \le x \le y \le 1. \qquad \square$$

We now dominate the right hand side of (2.5) to obtain the desired approximation.

<u>LEMMA 2.3</u> If $D(u_n)$ holds, $r \ge 1$ is any fixed integer, and if $n > (2Mr+1)k$, then, with the same notation as in Lemma 2.2,

$$\text{(2.6)} \qquad P\{M(I_1) \le u_n < M(I_1^*)\} \le \frac{1}{r} + 2r\,\alpha_{n,m}.$$

It then follows from Lemma 2.2 that

$$\text{(2.7)} \qquad P\{M_n \le u_n\} - P^k\{M_{[n/k]} \le u_n\} \to 0 \quad \text{as} \quad n \to \infty.$$

<u>PROOF</u> Since $n' > 2Mr+1$, we may choose intervals $E_1, \ldots, E_r$, each of length $m$, from $I_1 = \{1, 2, \ldots, n'-m\}$, $(n' = [n/k])$ so that they are separated from each other and from $I_1^*$ by at least $m$ $(k < m < n'$ again). Then

$$P\{M(I_1) \le u_n < M(I_1^*)\} \le P\left(\bigcap_{s=1}^{r} \{M(E_s) \le u_n\}, \{M(I_1^*) > u_n\}\right)$$

$$= P\left(\bigcap_{s=1}^{r} \{M(E_s) \le u_n\}\right) - P\left(\bigcap_{s=1}^{r} \{M(E_s) \le u_n\}, \{M(I_1^*) \le u_n\}\right).$$

By stationarity, $P\{M(E_s) \le u_n\} = P\{M(I_1^*) \le u_n\} = p$, say, and by Lemma 2.1, the two terms on the right differ from $p^r$ and $p^{r+1}$ (in absolute magnitude) by no more than $(r-1)\alpha_{n,m}$ and $r\alpha_{n,m}$, respectively. Hence

$$P\{M(I_1) \le u_n < M(I_1^*)\} \le p^r - p^{r+1} + 2r\,\alpha_{n,m},$$

from which (2.6) follows since $p^r - p^{r+1} \le 1/(r+1)$.

Finally by (2.5) and (2.6), taking $m = \ell_n$ according to (2.2), $(\ell_n = o(n))$

$$\limsup_{n\to\infty} |P\{M_n \le u_n\} - P^k\{M_{n'} \le u_n\}|$$

$$\le \frac{2k+1}{r} + (k + 2r(2k+1)) \limsup_{n\to\infty} \alpha_{n,\ell_n} = \frac{2k+1}{r},$$

from which it follows (by letting $r \to \infty$ on the right), that the left

hand side is zero. Thus (2.7) is proved. □

Gnedenko's Theorem now follows easily under general conditions.

THEOREM 2.4 Let $\{\xi_n\}$ be a stationary sequence, and $a_n > 0$ and $b_n$ given constants such that $P\{a_n(M_n - b_n) \leq x\}$ converges to a non-degenerate d.f. $G(x)$. Suppose that $D(u_n)$ is satisfied for $u_n = x/a_n + b_n$ for each real $x$. Then $G(x)$ has one of the three extreme value forms listed in Theorem 1.6.

PROOF Writing $u_n = x/a_n + b_n$ and using (2.7) (with $nk$ in place of $n$) we obtain (2.4). Hence if $F_n$ is the d.f. of $M_n$ then, as remarked in connection with (2.4), (1.11) holds for all $k$ since it holds by assumption for $k = 1$. Hence by Lemma 1.4, $G$ is max-stable and thus an extreme value type by Theorem 1.6. □

COROLLARY 2.5 The result remains true if the condition that $D(u_n)$ be satisfied for each $u_n = x/a_n + b_n$, is replaced by the requirement that Condition $D$ holds. (For then $D(u_n)$ is satisfied by any sequence, in particular by $u_n = x/a_n + b_n$ for each $x$.) □

## Convergence of $P\{M_n \leq u_n\}$ under dependence

The results so far have been concerned with the *possible* forms of limiting extreme value distributions. We now turn to the *existence* of such a limit, in that we formulate conditions under which (1.17) and (1.16) are equivalent for stationary sequences, i.e. conditions under which

(2.8) $\quad P\{M_n \leq u_n\} \to e^{-\tau}$

is equivalent to

(2.9) $\quad 1 - F(u_n) = P\{\xi_1 > u_n\} = \tau/n + o(1/n) \quad \text{as} \quad n \to \infty.$

We will show as a corollary that, under conditions of this type, $a_n(M_n - b_n)$ has the same limiting distribution as it would if the $\xi_n$ were i.i.d.

As may be seen from the derivation below, if (2.9) holds, then Condition $D(u_n)$ is sufficient to guarantee that $\liminf P\{M_n \le u_n\} \ge e^{-\tau}$. However, we need a further assumption to obtain the opposite inequality for the upper limit. Various forms of such an assumption may be used. Here we content ourselves with the following simple variant of conditions used in Watson (1954) and Loynes (1965); we refer to this as $D'(u_n)$.

<u>CONDITION $D'(u_n)$</u> will be said to hold for the stationary sequence $\{\xi_j\}$ and sequence $\{u_n\}$ of constants if

$$\text{(2.10)} \quad \limsup_{n\to\infty} n \sum_{j=2}^{[n/k]} P\{\xi_1 > u_n, \xi_j > u_n\} \to 0 \quad \text{as} \quad k \to \infty,$$

(where [ ] denotes the integer part).

Note that under (2.9), the level $u_n$ in (2.10) is such that there are on the average approximately $\tau$ exceedances of $u_n$ among $\xi_1, \ldots, \xi_n$, and thus $\tau/k$ among $\xi_1, \ldots, \xi_{[n/k]}$. The condition $D'(u_n)$ bounds the probability of more than one exceedance among $\xi_1, \ldots, \xi_{[n/k]}$. This will eventually ensure that there are no multiple points in the point process of exceedances, which of course is necessary in obtaining (as we later shall) a simple Poisson limit for this point process.

Our main result generalizes Theorem 1.10 to apply to stationary sequences under $D(u_n)$, $D'(u_n)$. A form of the "only if" part of this theorem was first proved by R. Davis (1978).

<u>THEOREM 2.6</u> If $D(u_n)$, $D'(u_n)$ hold for the stationary sequence $\{\xi_j\}$, then (2.8) and (2.9) are equivalent, i.e. $P\{M_n \le u_n\} \to e^{-\tau}$ if and only if $1 - F(u_n) \sim \tau/n$.

<u>PROOF</u> Fix $k$ and for each $n$ write $n' \ (= n'_{n,k}) = [n/k]$. Since $\{M_{n'} > u_n\} = \bigcup_{j=1}^{n'} \{\xi_j > u_n\}$ we have

$$\sum_{j=1}^{n'} P\{\xi_j > u_n\} - \sum_{1\le i<j\le n'} P\{\xi_i > u_n, \xi_j > u_n\} \le P\{M_{n'} > u_n\}$$
$$\le \sum_{j=1}^{n'} P\{\xi_j > u_n\}.$$

Using stationarity it follows simply that

(2.11) $\quad 1 - n'(1 - F(u_n)) \le P\{M_{n'} \le u_n\} \le 1 - n'(1 - F(u_n)) + S_n$

where $S_n = S_{n,k} = n' \sum_{j=2}^{n'} P\{\xi_1 > u_n, \xi_j > u_n\}$. Since $n' = [n/k]$, Condition $D'(u_n)$ gives $\limsup_{n\to\infty} S_n = k^{-1}o(1) = o(k^{-1})$ as $k \to \infty$.

Suppose now that (2.9) holds. Then $1 - F(u_n) \sim \tau/n \sim \tau/(n'k)$, so that $n \to \infty$ in (2.11) gives

$$1 - \tau/k \le \liminf_{n\to\infty} P\{M_{n'} \le u_n\} \le \limsup_{n\to\infty} P\{M_{n'} \le u_n\} \le 1 - \tau/k + o(1/k).$$

By taking the k - th power of each term and using (2.7) we have

$$(1 - \tau/k)^k \le \liminf_{n\to\infty} P\{M_n \le u_n\} \le \limsup_{n\to\infty} P\{M_n \le u_n\} \le (1 - \tau/k + o(1/k))^k.$$

Letting $k \to \infty$ we see that $\lim_{n\to\infty} P\{M_n \le u_n\}$ exists and equals $e^{-\tau}$, as required to show (2.8).

Conversely if (2.8) holds, i.e. $P\{M_n \le u_n\} \to e^{-\tau}$ as $n \to \infty$, we have from (2.11) (again with $n' = [n/k]$),

(2.12) $\quad 1 - P\{M_{n'} \le u_n\} \le n'(1 - F(u_n)) \le 1 - P\{M_{n'} \le u_n\} + S_n .$

But since $P\{M_n \le u_n\} \to e^{-\tau}$, (2.7) shows that $P\{M_{n'} \le u_n\} \to e^{-\tau/k}$, so that letting $n \to \infty$ in (2.12) we obtain, since $n' \sim n/k$,

$$1 - e^{-\tau/k} \le \frac{1}{k} \liminf_{n\to\infty} n(1 - F(u_n)) \le \frac{1}{k} \limsup_{n\to\infty} n(1 - F(u_n))$$
$$\le 1 - e^{-\tau/k} + o(1/k)$$

from which (multiplying by $k$ and letting $k \to \infty$) we see that $n(1 - F(u_n)) \to \tau$ so that (2.9) holds. □

We remark here that if $D(u_n)$, $D'(u_n)$ hold and it is assumed that (2.9) holds just for some subsequence $\{n_j\}$ of integers, i.e. $1 - F(u_{n_j}) \sim \tau/n_j$ as $j \to \infty$, then the same proof shows that (2.8) holds for that

subsequence, i.e. $P\{M_{n_j} \le u_{n_j}\} \to e^{-\tau}$. This observation may be used to give an alternative proof of the statement that (2.8) implies (2.9) in this theorem, by assuming the existence of some subsequence $n_j$ for which $n_j(1-F(u_{n_j})) \to \tau' \neq \tau$, $0 \le \tau' \le \infty$.

As a corollary of Theorem 2.6 we may obtain a limiting result for the maximum $M(I_n)$ of $\xi_j$ for $j$ in any "interval" whose length is asymptotically proportional to $n$.

COROLLARY 2.7 Let $\{\xi_n\}$ be a stationary sequence, and $\{u_n(\tau)\}$, $\{u_n(\theta\tau)\}$ real sequences satisfying (2.9), and (2.9) with $\theta\tau$ in place of $\tau$, respectively, i.e.

(2.13) $1 - F(u_n(\theta\tau)) \sim \theta\tau/n$ as $n \to \infty$,

where $\theta > 0$, $\tau \ge 0$. Suppose that $D(u_n(\theta\tau))$, $D'(u_n(\theta\tau))$ hold. Then if $\{I_n\}$ is a sequence of intervals of integers with $\nu_n$ members where $\nu_n \sim \theta n$, we have

(2.14) $P\{M(I_n) \le u_n(\tau)\} \to e^{-\theta\tau}$ as $n \to \infty$.

PROOF By stationarity it is sufficient to show that $P\{M_{\nu_n} \le u_n(\tau)\} \to e^{-\theta\tau}$. Now it follows at once from Theorem 2.6 that $P\{M_{\nu_n} \le u_{\nu_n}(\theta\tau)\} \to e^{-\theta\tau}$ and hence it is only necessary to show that

(2.15) $P\{M_{\nu_n} \le u_n(\tau)\} - P\{M_{\nu_n} \le u_{\nu_n}(\theta\tau)\} \to 0$.

If $u_n(\tau) > u_{\nu_n}(\theta\tau)$ the left hand side of (2.15) is

$$P\{u_{\nu_n}(\theta\tau) < M_{\nu_n} \le u_n(\tau)\} \le P\left(\bigcup_{j=1}^{\nu_n} \{u_{\nu_n}(\theta\tau) < \xi_j \le u_n(\tau)\}\right)$$

$$\le \nu_n\{F(u_n(\tau)) - F(u_{\nu_n}(\theta\tau))\}.$$

If $u_n(\tau) < u_{\nu_n}(\theta\tau)$ the same result holds with reversed signs so that the left hand side of (2.15) is always dominated in modulus by

$$\nu_n|F(u_n(\tau)) - F(u_{\nu_n}(\theta\tau))| = \nu_n|1 - F(u_{\nu_n}(\theta\tau)) - (1 - F(u_n(\tau)))|$$

$$= \nu_n\left|\frac{\theta\tau}{\nu_n}(1 + o(1)) - \frac{\tau}{n}(1 + o(1))\right|$$

$$= o(1) \quad \text{as} \quad n \to \infty$$

as required. □

Note that if $u_n(\theta\tau)$ is defined to satisfy (2.13) we may obtain $u_n(\tau)$ satisfying (2.9) by taking $u_n(\tau) = u_{[n\theta]}(\theta\tau)$, since, by (2.13),

$$1 - F(u_{[n\theta]}(\theta\tau)) \sim \theta\tau/[n\theta] \sim \tau/n.$$

(Of course the corollary applies with any other $u_n(\tau)$ satisfying (2.8) as well as this one.)

Putting this slightly differently, if $u_n(\tau)$ is chosen to satisfy (2.9), we may define $u_n(\theta\tau)$ to satisfy (2.13) by writing $u_n(\theta\tau) = u_{[n/\theta]}(\tau)$. That is, given $u_n(\tau)$ satisfying (2.9) for some $\tau$, we may define $u_n(\tau')$ satisfying (2.9) for all $\tau' > 0$ by taking $\theta = \tau'/\tau$. It is also useful to know that when $u_n(\tau')$ is thus defined, if $D(u_n(\tau))$ or $D'(u_n(\tau))$ holds, then so does $D(u_n(\tau'))$, or $D'(u_n(\tau'))$, for $\tau' < \tau$. We show this in the following lemma.

<u>LEMMA 2.8</u> Let $u_n(\tau)$ be defined, satisfying (2.9) for a fixed $\tau > 0$. For $\theta > 0$ define

(2.16) $u_n(\theta\tau) = u_{[n/\theta]}(\tau)$.

Then

(i) $u_n(\theta\tau)$ satisfies (2.13),

(ii) if $\theta < 1$, and $D(u_n)$ holds with $u_n = u_n(\tau)$, it holds also for $u_n(\theta\tau)$,

(iii) the same is true for $D'(u_n)$ instead of $D(u_n)$.

<u>PROOF</u> (i) By (2.9),

$$1-F(u_{[n/\theta]}(\tau)) \sim \tau/[n/\theta] \sim \theta\tau/n.$$

(ii) If $D(u_n(\tau))$ holds and $1 \leq i_1 < \ldots < i_p < j_1, \ldots < j_q \leq n$, $j_1 - i_p \geq \ell$, we have, writing $\underline{i} = (i_1, \ldots, i_p)$, $\underline{j} = (j_1, \ldots, j_q)$ for brevity,

$$|F_{\underline{i},\underline{j}}(u_n(\theta\tau)) - F_{\underline{i}}(u_n(\theta\tau))F_{\underline{j}}(u_n(\theta\tau))|$$

$$= |F_{\underline{i},\underline{j}}(u_{[n/\theta]}(\tau)) - F_{\underline{i}}(u_{[n/\theta]}(\tau))F_{\underline{j}}(u_{[n/\theta]}(\tau))|$$

which does not exceed $\alpha'_{n,\ell} = \alpha_{[n/\theta],\ell}$ (since $j_q \leq n \leq [n/\theta]$). If $\alpha_{n,\ell_n} \to 0$ then $\alpha'_{n,\ell'_n} \to 0$ with $\ell'_n = \ell_{[n/\theta]} = o(n)$, so that (ii) follows.

(iii) also follows simply since

$$n \sum_{j=2}^{[n/k]} P\{\xi_1 > u_n(\theta\tau),\ \xi_j > u_n(\theta\tau)\}$$

$$= n \sum_{j=2}^{[n/k]} P\{\xi_1 > u_{[n/\theta]}(\tau),\ \xi_j > u_{[n/\theta]}(\tau)\}$$

$$\leq [n/\theta] \sum_{j=2}^{[[n/\theta]/k]} P\{\xi_1 > u_{[n/\theta]}(\tau),\ \xi_j > u_{[n/\theta]}(\tau)\}.$$

By $D'(u_n)$, the upper limit of this expression over $n$ (or $[n/\theta]$) tends to zero as $k \to \infty$, so that (iii) follows. □

From these results we may see simply that, under $D$ and $D'$ conditions the asymptotic distributions of $M(I_n)$ is of the same type as that of $M_n$, when $I_n$ has $\nu_n \sim \theta n$ members.

<u>THEOREM 2.9</u> Let $\{\xi_n\}$ be a stationary sequence, let $a_n > 0$, $b_n$ be constants, and suppose that

$$P\{a_n(M_n - b_n) \leq x\} \to G(x) \quad \text{as} \quad n \to \infty,$$

for a non-degenerate d.f. $G$. Suppose $D(u_n)$, $D'(u_n)$ hold for all $u_n$ of the form $x/a_n + b_n$, and let $I_n$ be an interval containing $\nu_n \sim \theta n$ integers for some $\theta$, $0 < \theta \leq 1$. Then

$$P\{a_n(M(I_n) - b_n) \leq x\} \to G^\theta(x) \quad \text{as} \quad n \to \infty.$$

PROOF Consider a point $x$ with $G(x) > 0$ and write $\tau = -\log G(x)$, $u_n = u_n(\tau) = x/a_n + b_n$. By Theorem 2.6 we see that (2.9) holds. Then by Lemma 2.8 it follows that $D(u_n(\theta\tau))$, $D'(u_n(\theta\tau))$ hold with $u_n(\theta\tau)$ defined by (2.16), and (2.13) also holds. Hence, by Corollary 2.7, $P\{M(I_n) \leq u_n\} \to e^{-\theta\tau}$, which gives the desired result on writing $u_n = x/a_n + b_n$, $\tau = -\log G(x)$. □

Note that since $G$ is an extreme value d.f., it is max-stable, and Corollary 1.5 shows that $G^\theta$ is of the same type as $G$, i.e. $G^\theta(x) = G(ax+b)$ for some $a > 0$, $b$. Also we may clearly modify the statement of the theorem slightly to include $\theta > 1$, or to compare $M(I_n)$ with $M(I_n')$ for some other interval $I_n'$ containing $\nu_n' \sim \theta' n$ integers.

Associated sequence of independent variables

We now write $\hat{M}_n$ for the maximum of $n$ i.i.d. random variables with the same marginal d.f. $F$ as each $\xi_n$; following Loynes (1965) we may call these the "independent sequence associated with $\{\xi_n\}$". The following result is then a corollary of Theorem 2.6.

THEOREM 2.10 Let $D(u_n)$, $D'(u_n)$ be satisfied for the stationary sequence $\{\xi_n\}$. Then $P\{M_n \leq u_n\} \to \theta > 0$ if and only if $P\{\hat{M}_n \leq u_n\} \to \theta$.

PROOF The condition $P\{\hat{M}_n \leq u_n\} \to \theta$ may be rewritten as $P\{\hat{M}_n \leq u_n\} \to e^{-\tau}$ with $\tau = -\log\theta$, and, by Theorem 1.10, holds if and only if $1 - F(u_n) \sim \tau/n$. By Theorem 2.6 the same is true for $P\{M_n \leq u_n\}$ so that the result follows. □

We may also deduce at once that the limiting distribution of $a_n(M_n - b_n)$ is the same as that which would apply if the $\xi_n$ were i.i.d., i.e. it is the same as that of $a_n(\hat{M}_n - b_n)$, under conditions $D(u_n)$ and $D'(u_n)$. Part of this result was proved in Loynes (1965) under conditions which include strong mixing.

<u>THEOREM 2.11</u> Suppose that $D(u_n)$, $D'(u_n)$ are satisfied for the stationary sequence $\{\xi_n\}$ when $u_n = x/a_n + b_n$, for each $x$, ($\{a_n > 0\}$, $\{b_n\}$ being given sequences of constants.) Then $P\{a_n(M_n - b_n) \le x\} \to G(x)$ for some non-degenerate $G$ if and only if $P\{a_n(\hat{M}_n - b_n) \le x\} \to G(x)$.

<u>PROOF</u> Under either assumption $G$ is an extreme value distribution by Gnedenko's Theorem, and hence continuous. If $G(x) > 0$, the equivalence follows from Theorem 2.10 with $\theta = G(x)$.

On the other hand, if e.g. $P\{a_n(M_n - b_n) \le x\} \to G(x)$ for all $x$, and $G(x_0) = 0$ we have $P\{a_n(\hat{M}_n - b_n) \le x_0\} \le P\{a_n(\hat{M}_n - b_n) \le y\} \to G(y)$ for any $y$ with $G(y) > 0$. By letting $y$ decrease we see that $\limsup_{n\to\infty} P\{a_n(\hat{M}_n - b_n) \le x_0\}$ is zero, so that $P\{a_n(\hat{M}_n - b_n) \le x_0\} \to 0$. This, and the corresponding result obtained by interchanging $M_n$ and $\hat{M}_n$, complete the proof. □

## CHAPTER 3

## THE NORMAL CASE

In this chapter we consider a stationary normal sequence $\{\xi_n\}$ with zero means, unit variances, and covariances $r_n = E(\xi_j \xi_{j+n})$. As noted in Chapter 2, Berman (1964) has given simple conditions on $r_n$ to ensure that (1.19) holds, i.e. that $M_n = \max(\xi_1, \ldots, \xi_n)$ has a limiting distribution of the double exponential type,

$$P\{a_n(M_n - b_n) \leq x\} \to \exp(-e^{-x}), \quad n \to \infty,$$

with

$$a_n = (2 \log n)^{1/2}$$

$$b_n = a_n - (2a_n)^{-1}\{\log \log n + \log 4\pi\}.$$

One of Berman's results is that it suffices that

$$(3.1) \qquad r_n \log n \to 0 \quad \text{as} \quad n \to \infty,$$

and we will devote the first part of this chapter for a proof of this. However, (3.1) can be replaced by an appropriate Cesàro convergence,

$$\frac{1}{n} \sum_{k=1}^{n} |r_k| \log k \, e^{\gamma |r_k| \log k} \to 0 \quad \text{as} \quad n \to \infty,$$

for some $\gamma > 2$, and this and some other conditions will be discussed at the end of the chapter.

It has been shown by Mittal and Ylvisaker (1975) that (3.1), and therefore also the Cesàro convergence, is rather close to being necessary, in that if $r_n \log n \to \gamma > 0$ a different limit applies. In fact, this strong dependence destroys the asymptotic independence of extremes in disjoint intervals (cf. Lemmas 2.1-2.3). We return to these matters in more detail in Chapter 5.

In this section we use the general results from Chapter 2 to derive the double exponential law for normal sequences under some different covariance assumptions, starting with the transparent condition (3.1)

and proving that this implies $D(u_n)$ and $D'(u_n)$. (This approach does not reduce the amount of work required, but does bring the normal case within the general framework.)

Throughout this and all subsequent chapters $\Phi$, $\phi$, will denote the standard normal distribution function and density respectively. We shall repeatedly have occasion to use the well known relation for the tail of $\Phi$, stated here for easy reference

(3.2) $\quad 1-\Phi(u) \sim \phi(u)/u \quad$ as $\quad u \to \infty.$

<u>Double exponential limit if $r_n \log n \to 0$</u>

Our aim is to show that if (3.1) holds, then $D(u_n)$ and $D'(u_n)$ are satisfied when

(3.3) $\quad 1-\Phi(u_n) \sim \tau/n.$

We can then conclude that (2.8) holds and obtain (1.19) for this sequence.

First, it will be convenient to prove a "technical" lemma (given in Berman (1964)), which is of a type that is very useful in both discrete and continuous cases. The proof is elementary, though some slightly messy calculation is involved.

Before stating the lemma, we note the easily proved fact that if $r_n \to 0$ as $n \to \infty$ then $|r_n|$ cannot equal 1 for any $n \neq 0$. (For if $|r_n| = 1$ for some $n \neq 0$, it follows (by normality) that $\xi_1$ and $\xi_{n+1}$ are linearly related, as are also $\xi_{n+1}$ and $\xi_{2n+1}$, and hence so are $\xi_1$, $\xi_{2n+1}$, so that $|r_{2n}| = 1$. In this way it follows that $|r_{kn}| = 1$ for all $k$ contradicting the requirement that $r_n \to 0$). Hence it is easy to see that $|r_n|$ is actually bounded away from 1 for all $n \geq 1$, i.e. $\sup_{n\geq 1} |r_n| = \delta < 1$.

<u>LEMMA 3.1</u> Suppose $\{r_n\}$ satisfies (3.1). Then

(3.4) $\quad n \sum_{j=1}^{n} |r_j| e^{-u_n^2/(1+|r_j|)} \to 0 \quad$ as $\quad n \to \infty$

if $1-\Phi(u_n) \sim \tau/n$ for some constant $\tau$.

<u>PROOF</u> By using (3.2) and (3.3) we see that

$$(3.5) \qquad \text{(i)} \quad e^{-u_n^2/2} \sim K\, u_n/n$$

$$\text{(ii)} \quad u_n \sim (2 \log n)^{1/2}$$

using $K$ as a constant whose value may change from line to line. As above, let $\delta = \sup_{n \geq 1} |r_n| < 1$, and let $\alpha$ be a constant such that $0 < \alpha < \frac{1-\delta}{1+\delta}$.

Split the sum in (3.4) into two parts, the first for $1 \leq j \leq [n^\alpha]$ and the second for $[n^\alpha] < j \leq n$. The first sum is dominated by

$$n\, n^\alpha e^{-u_n^2/(1+\delta)} = n^{1+\alpha} (e^{-u_n^2/2})^{2/(1+\delta)} \leq K\, n^{1+\alpha} (u_n/n)^{2/(1+\delta)}$$

$$\leq K\, n^{1+\alpha-2/(1+\delta)} (\log n)^{1/(1+\delta)}$$

(where (3.5), (i) and (ii), have been used). This tends to zero since $1 + \alpha - \frac{2}{1+\delta} < 0$ from the choice of $\alpha$.

To deal with the second part we define

$$\delta_n = \sup_{m \geq n} |r_m|$$

and note that

$$\delta_n \log n \leq \sup_{m \geq n} |r_m| \log m \to 0 \quad \text{as} \quad n \to \infty.$$

Now writing $p = [n^\alpha]$ we have for the second part of (3.4),

$$n \sum_{j=p+1}^{n} |r_j| e^{-u_n^2/(1+|r_j|)} \leq n\, \delta_p\, e^{-u_n^2} \sum_{j=p+1}^{n} e^{u_n^2 |r_j|/(1+|r_j|)}$$

$$\leq n^2 \delta_p e^{-u_n^2} e^{u_n^2 \delta_p} \leq K\, \delta_p u_n^2 e^{\delta_p u_n^2}$$

by (3.5), (i). But by (3.5), (ii),

$$\delta_p u_n^2 \sim 2\, \delta_{n^\alpha} \log n = \frac{2}{\alpha}\, \delta_{n^\alpha} \log n^\alpha,$$

which tends to zero. Thus the exponential term above tends to one and the remaining product tends to zero, so that the desired result follows. □

We now attack the main task, using a very instructive method developed, in various ways, by Slepian (1962), Berman (1964, 1971a), and

Cramér (see Cramér and Leadbetter (1967)). The conditions $D(u_n)$ and $D'(u_n)$ will both follow from the following lemma. It is stated in some generality, even though needed here only in a simple special form – the complete formulation being used in later chapters.

LEMMA 3.2 Suppose $\xi_1, \ldots, \xi_n$ are standard normal variables with covariance matrix $\Lambda^1 = (\Lambda^1_{ij})$, and $\eta_1, \ldots, \eta_n$ similarly with covariance matrix $\Lambda^0 = (\Lambda^0_{ij})$, and let $\rho_{ij} = \max(|\Lambda^1_{ij}|, |\Lambda^0_{ij}|)$. Further, let $\underline{u} = (u_1, \ldots, u_n)$ be a vector of real numbers and write $u = \min(|u_1|, \ldots, |u_n|)$. Then

$$(3.6)\qquad P\{\xi_j \le u_j \text{ for } j = 1, \ldots, n\} - P\{\eta_j \le u_j \text{ for } j = 1, \ldots, n\}$$
$$\le \frac{1}{2\pi} \sum_{1\le i<j\le n} (\Lambda^1_{ij} - \Lambda^0_{ij})^+ (1-\rho^2_{ij})^{-1/2} e^{-u^2/(1+\rho_{ij})}$$

where $(x)^+ = \max(0,x)$.

In particular, if $\eta_1, \ldots, \eta_n$ are independent and $\delta = \max_{i\neq j} |\Lambda^1_{ij}| < 1$, then for any real $u$ and integers $1 \le \ell_1 < \ldots < \ell_s$,

$$|P\{\xi_{\ell_j} \le u \text{ for } j = 1, \ldots, s\} - \Phi(u)^s|$$
$$\le K \sum_{1\le i<j\le s} |r_{ij}| e^{-u^2/(1+|r_{ij}|)}$$

where $r_{ij} = \Lambda^1_{\ell_i \ell_j}$ is the correlation between $\xi_{\ell_i}$ and $\xi_{\ell_j}$, and $K$ is a constant (depending on $\delta$).

If furthermore $\{\xi_n\}$ is stationary with $r_i = \mathrm{Cov}(\xi_1, \xi_{1+i})$, $1 \le \ell_1 < \ldots < \ell_s \le n$,

$$(3.7)\qquad |P\{\xi_{\ell_j} \le u \text{ for } i = 1, \ldots, s\} - \Phi(u)^s| \le K n \sum_{i=1}^{n} |r_i| e^{-u^2/(1+|r_i|)} .$$

PROOF We shall suppose that $\Lambda^1$ and $\Lambda^0$ are positive *definite* (as opposed to *semi*-definite) and hence that $(\xi_1, \ldots, \xi_n)$ and $(\eta_1, \ldots, \eta_n)$ have joint *densities* $f_1$ and $f_0$, respectively. (The semi-definite case is easily dealt with by considering $\xi_i + \varepsilon_i$ and $\eta_i + \varepsilon_i$, where the $\varepsilon_i$ are independent normal variables with mean zero and then letting $\mathrm{Var}(\varepsilon_i) \to 0$, using continuity.) Clearly

$$P\{\xi_j \leq u_j \text{ for } j = 1, \ldots, n\} = \int \overset{u}{\underset{-\infty}{\cdots}} \int f_1(y_1, \ldots, y_n) d\underline{y}$$

$$P\{\eta_j \leq u_j \text{ for } j = 1, \ldots, n\} = \int \overset{u}{\underset{-\infty}{\cdots}} \int f_0(y_1, \ldots, y_n) d\underline{y}$$

where $f_1$, $f_0$ are the normal density functions based on the covariance matrices $\Lambda^1$, $\Lambda^0$, the integration ranges being $\{\underline{y};\ y_j \leq u_j,\ j = 1, \ldots, n\}$.

If we write $\Lambda_h = h\,\Lambda^1 + (1-h)\Lambda^0$, $(0 \leq h \leq 1)$, the matrix $\Lambda_h$ is positive definite with units down the main diagonal and elements $h\,\Lambda^1_{ij} + (1-h)\Lambda^0_{ij}$ for $i \neq j$. Let $f_h$ be the n-dimensional normal density based on $\Lambda_h$, and

$$F(h) = \int \overset{u}{\underset{-\infty}{\cdots}} \int f_h(y_1, \ldots, y_n) d\underline{y}.$$

The left hand side of (3.6) is then easily recognized as $F(1) - F(0)$. Now

$$F(1) - F(0) = \int_0^1 F'(h)dh$$

where

$$F'(h) = \int \overset{u}{\underset{-\infty}{\cdots}} \int \frac{\partial f_h(y_1, \ldots, y_n)}{\partial h} d\underline{y}.$$

The density $f_h$ depends on $h$ only through the elements $\Lambda^h_{ij}$ of $\Lambda_h$ (regarding $f_h$ as a function of $\Lambda^h_{ij}$ for $i \leq j$, say). We have $\Lambda^h_{ii} = 1$ independent of $h$, while for $i < j$, $\Lambda^h_{ij} = h\,\Lambda^1_{ij} + (1-h)\Lambda^0_{ij}$ so that $\partial\Lambda^h_{ij}/\partial h = \Lambda^1_{ij} - \Lambda^0_{ij}$. Thus

$$F'(h) = \sum_{i \leq j} \int \overset{u}{\underset{-\infty}{\cdots}} \int \frac{\partial f_h}{\partial \Lambda^h_{ij}} \cdot \frac{\partial \Lambda^h_{ij}}{\partial h} d\underline{y} = \sum_{i<j} (\Lambda^1_{ij} - \Lambda^0_{ij}) \int \overset{u}{\underset{-\infty}{\cdots}} \int \frac{\partial f_h}{\partial \Lambda^h_{ij}} d\underline{y}.$$

Now a useful property of the multidimensional normal density is that its derivative with respect to a covariance $\Lambda_{ij}$ is the same as the second mixed derivative with respect to the corresponding variables $y_i$, $y_j$, (cf. Cramér and Leadbetter (1967), p. 26), i.e.

$$\frac{\partial f_h}{\partial \Lambda_{ij}} = \frac{\partial^2 f_h}{\partial y_i \partial y_j}.$$

Thus

$$F'(h) = \sum_{i<j} (\Lambda^1_{ij} - \Lambda^0_{ij}) \int \overset{u}{\underset{-\infty}{\cdots}} \int \frac{\partial^2 f_h}{\partial y_i \partial y_j} d\underline{y}.$$

The $y_i$ and $y_j$ integrations may be done at once to give

$$(3.8) \qquad \sum_{i<j} (\Lambda^1_{ij} - \Lambda^0_{ij}) \int \overset{\underline{u}'}{\underset{-\infty}{\cdots}} \int f_h(y_i = u_i, y_j = u_j)\, d\underline{y}'$$

where $f_h(y_i = u_i, y_j = u_j)$ denotes the function of $n-2$ variables formed by putting $y_i = u_i$, $y_j = u_j$, the integration being over the remaining variables.

Further, we can dominate the last integral by letting the variables run from $-\infty$ to $+\infty$. But

$$\int \overset{\infty}{\underset{-\infty}{\cdots}} \int f_h(y_i = u_i, y_j = u_j)\, d\underline{y}'$$

is just the bivariate density, evaluated at $(u_i, u_j)$, of two standard normal random variables with correlation $\Lambda^h_{ij}$, and may therefore be written

$$\frac{1}{2\pi(1 - (\Lambda^h_{ij})^2)^{1/2}} \exp\left\{- \frac{1}{2(1 - (\Lambda^h_{ij})^2)} (u_i^2 - 2\Lambda^h_{ij} u_i u_j + u_j^2)\right\}.$$

Now, since $|\Lambda^h_{ij}| = |h\Lambda^1_{ij} + (1-h)\Lambda^0_{ij}| \le \max(|\Lambda^1_{ij}|, |\Lambda^0_{ij}|) = \rho_{ij}$, and $u = \min(|u_1|, \ldots, |u_n|) \le \min(|u_i|, |u_j|)$, it may be easily shown that the above expression is not greater than

$$\frac{1}{2\pi(1 - \rho^2_{ij})^{1/2}} \exp\{-u^2/(1 + \rho_{ij})\}.$$

(Note that if $Q_\rho(x, y) = x^2 - 2\rho x y + y^2$ and $0 \le x < y$ then $Q_\rho(x, y) \ge Q_\rho(x, x)$ and in general $Q_\rho(x, y) \ge Q_{|\rho|}(|x|, |y|)$.) Eliminating the possible negative terms in (3.8), we thus obtain

$$F'(h) \le \frac{1}{2\pi} \sum_{i<j} (\Lambda^1_{ij} - \Lambda^0_{ij})^+ (1 - \rho^2_{ij})^{-1/2} e^{-u^2/(1+\rho_{ij})},$$

and since $F(1) - F(0) = \int_0^1 F'(h)\, dh$, this proves the lemma. □

The conditions $D(u_n)$ and $D'(u_n)$ now follow simply.

LEMMA 3.3 Suppose that $1-\Phi(u_n) \sim \tau/n$ and that (3.4) holds with $\delta = \sup_{n\geq 1}|r_n| < 1$ (which will be so, in particular, if (3.1) is satisfied). Then both $D(u_n)$ and $D'(u_n)$ hold.

PROOF From (3.7) with $s = 2$, $\ell_1 = 1$, $\ell_2 = j$, we have

$$|P\{\xi_1 \leq u_n, \xi_j \leq u_n\} - \Phi^2(u_n)| \leq K|r_{j-1}|e^{-u_n^2/(1+|r_{j-1}|)},$$

whence, by simple manipulation, ($\xi_1$ and $\xi_2$ being each standard normal),

$$|P\{\xi_1 > u_n, \xi_j > u_n\} - (1-\Phi(u_n))^2| \leq K|r_{j-1}|e^{-u_n^2/(1+|r_{j-1}|)}.$$

Thus

$$n\sum_{j=2}^{[n/k]} P\{\xi_1 > u_n, \xi_j > u_n\} \leq \frac{\tau^2}{k} + Kn\sum_{j=1}^{n}|r_j|e^{-u_n^2/(1+|r_j|)} + o(1),$$

from which $D'(u_n)$ follows by (3.4).

It follows also from (3.7) that if $1 \leq \ell_1 \leq \cdots \leq \ell_s \leq n$, then

$$|F_{\ell_1\ldots\ell_s}(u_n) - \Phi^s(u_n)| \leq Kn\sum_{j=1}^{n}|r_j|e^{-u_n^2/(1+|r_j|)}.$$

Suppose now that $1 \leq i_1 < \cdots < i_p < j_1 < \cdots < j_q \leq n$. Identifying $\{\ell_1, \ldots, \ell_s\}$ in turn with $\{i_1, \ldots, i_p, j_1, \ldots, j_q\}$, $\{i_1, \ldots, i_p\}$ and $\{j_1, \ldots, j_q\}$ we thus have

$$|F_{i_1\ldots i_p,j_1\ldots j_q}(u_n) - F_{i_1\ldots i_p}(u_n)\,F_{j_1\ldots j_q}(u_n)|$$
$$\leq 3\,Kn\sum_{j=1}^{n}|r_j|e^{-u_n^2/(1+|r_j|)},$$

which tends to zero by (3.4). Thus $D(u_n)$ is satisfied. (In fact $\lim_{n\to\infty}\alpha_{n,\ell} = 0$ for each $\ell$.) □

Our main results now follow from Theorem 2.6 and Lemma 3.3.

THEOREM 3.4 If $\{\xi_n\}$ is a stationary normal sequence satisfying (3.1), then $P\{M_n \leq u_n\} \to e^{-\tau}$ if and only if $1-\Phi(u_n) \sim \tau/n$. □

THEOREM 3.5 If $\{\xi_n\}$ is a stationary normal sequence satisfying (3.1),

then (1.19) holds, i.e.

$$P\{a_n(M_n - b_n) \le x\} \to \exp(-e^{-x}) \quad \text{as} \quad n \to \infty$$

where

$$a_n = (2 \log n)^{1/2}$$

$$b_n = (2 \log n)^{1/2} - \frac{1}{2}(2 \log n)^{-1/2}\{\log \log n + \log 4\pi\}.$$

PROOF Since by Theorem 1.11 the result holds for the associated independent sequence, if we write $u_n = x/a_n + b_n$, we have by Theorem 1.10 that $1 - \Phi(u_n) \sim \tau/n$ where $\tau = e^{-x}$. Thus by Lemma 3.3 both $D(u_n)$ and $D'(u_n)$ hold and it follows from Theorem 2.6 that $P\{M_n \le u_n\} \to e^{-\tau}$. Rephrased, this is precisely the desired conclusion of the theorem. □

Weaker dependence assumptions for double exponential limits

The condition (3.1) that $r_n \log n \to 0$ is rather close to what is necessary for the maximum of a stationary normal sequence to behave like that of the associated independent sequence.

As we have seen, it is the convergence (3.4) which makes it possible to prove $D(u_n)$ and $D'(u_n)$, and one could therefore be tempted to use (3.4) as an indeed very weak condition. Since it is not very transparent, depending as it does on the level $u_n$, other conditions, which also restrict the size of $r_n$ for large $n$, have been proposed occasionally. In Berman (1964) it is shown that (3.1) can be replaced by

$$(3.9) \qquad \sum_{n=0}^{\infty} r_n^2 < \infty$$

which is a special case of

$$(3.10) \qquad \sum_{n=0}^{\infty} r_n^p < \infty, \text{ for some } p > 0.$$

There is no implication between (3.1) and condition (3.10), but they both imply the following weak condition,

$$(3.11)\quad \frac{1}{n}\sum_{k=1}^{n}|r_k|\log k\, e^{\gamma|r_k|\log k} \to 0 \quad \text{as} \quad n \to \infty$$

for some $\gamma > 2$, as is proved in Leadbetter et al. (1979) and below after Theorem 3.7. We now show that (3.11) implies the relevant $D(u_n)$, $D'(u_n)$ conditions.

<u>LEMMA 3.6</u> If $r_n \to 0$ as $n \to \infty$, $\{r_n\}$ satisfies (3.11), and if $1-\Phi(u_n) \sim \tau/n$, then (3.4) holds, and consequently (by Lemma 3.3) also $D(u_n)$ and $D'(u_n)$ hold.

<u>PROOF</u> Using the notation in the proof of Lemma 3.1, let $\delta = \sup_{n\geq 1}|r_n| < 1$, take $\beta = 2/\gamma$ and let $\alpha$ be a constant such that $0 < \alpha < \min(\beta, \frac{1-\delta}{1+\delta})$.

Split the sum in (3.4) into three parts, the first for $1 \leq j \leq [n^\alpha]$, the second for $[n^\alpha] < j \leq [n^\beta]$ and the third for $[n^\beta] < j \leq n$. The first sum tends to zero as in Lemma 3.1.

Writing $\delta_n = \sup_{m\geq n}|r_m|$, $p = [n^\alpha]$, and $q = [n^\beta]$ and using (3.5), i.e.

$$e^{-u_n^2/2} \sim Ku_n/n \sim K(2\log n)^{1/2}/n,$$

we have for the second part of (3.4),

$$n\sum_{k=p+1}^{q}|r_k|e^{-u_n^2/(1+|r_k|)} \leq n^{1+\beta}e^{-u_n^2}e^{u_n^2\delta_p} \leq Kn^{\beta-1}u_n^2\, e^{\delta_p u_n^2}$$
$$\leq Kn^{\beta-1}u_n^2\, n^{2\delta_p}$$

which obviously tends to zero, since $\beta < 1$.

Finally, for the last part of (3.4) we have, again using (3.5),

$$n\sum_{k=q+1}^{n}|r_k|e^{-u_n^2/(1+|r_k|)} \leq Kn\sum_{k=q+1}^{n}|r_k|(u_n/n)^{2/(1+|r_k|)}$$
$$\leq Kn^{-1}\log n\sum_{k=q+1}^{n}|r_k|e^{2|r_k|\log n}.$$

For $k > q$ we have $\log k \geq \beta\log n$, and hence this expression is not larger than

$$Kn^{-1}\sum_{k=q+1}^{n}|r_k|\log k\, e^{2\beta^{-1}|r_k|\log k} \leq Kn^{-1}\sum_{k=1}^{n}|r_k|\log k\, e^{\gamma|r_k|\log k}.$$

By (3.11) this tends to zero as $n \to \infty$, which concludes the proof of (3.4). □

The next result (extending Theorem 3.5) now follows by exactly the same proof as Theorem 3.5.

THEOREM 3.7 If $\{\xi_n\}$ is a stationary normal sequence with $r_n \to 0$ satisfying (3.11), the distribution of the normalized maxima has a limiting distribution of Type I with $a_n$ and $b_n$ as in Theorem 3.5. □

A few remarks may be in order here to elucidate the content of condition (3.11).

Define for each positive $x$, the set $\theta_n(x) = \{k;\ 1 \le k \le n,\ |r_k| \log k > x\}$ and let $\nu_n(x)$ be the number of elements in $\theta_n(x)$. Consider the following condition (which we shall see is slightly stronger than (3.11)),

(3.12)' $n^{-1} \sum_{k=1}^{n} |r_k| \log k \to 0$, as $n \to \infty$, and

$$\nu_n(K) = O(n^\eta) \quad \text{for some} \quad K > 0,\ \eta < 1,$$

and the equivalent condition

(3.12)" $\nu_n(\varepsilon) = o(n)$ for all $\varepsilon > 0$, and

$$\nu_n(K) = O(n^\eta) \quad \text{for some} \quad K > 0,\ \eta < 1.$$

Obviously (3.1) implies (3.12)'. Further, if

$$\sum_{k=1}^{\infty} |r_k|^p < \infty$$

for some $p > 0$ then, since

$$\sum_{k=1}^{\infty} |r_k|^p \ge \sum_{k \in \theta_n(x)} |r_k|^p \ge \nu_n(x)(x/\log n)^p,$$

it follows that $\nu_n(x) = O((\log n)^p)$. In particular, taking $p = 2$ we see that also (3.9) and (3.10) imply (3.12)", so that both (3.1) and (3.10) are stronger than (3.12)' and (3.12)". The following lemma states that (3.12)' or (3.12)" imply (3.11) and consequently that both (3.1) and (3.10) imply (3.11).

LEMMA 3.8 If $r_n \to 0$ as $n \to \infty$, then (3.12)' and (3.12)" both imply (3.11).

PROOF It is easily seen that (3.12)' and (3.12)" are equivalent so we need only show that (3.12)' implies (3.11). We have

$$(3.13)\quad n^{-1}\sum_{k=1}^{n}|r_k|\log k\, e^{\gamma|r_k|\log k} = n^{-1}\sum_{\substack{k=1\\ k\notin\theta_n(K)}}^{n}|r_k|\log k\, e^{\gamma|r_k|\log k} + n^{-1}\sum_{k\in\theta_n(K)}|r_k|\log k\, e^{\gamma|r_k|\log k},$$

and proceed to estimate the sums on the right separately, assuming that (3.12)' holds. Now

$$n^{-1}\sum_{\substack{k=1\\ k\notin\theta_n(K)}}^{n}|r_k|\log k\, e^{\gamma|r_k|\log k} \le e^{\gamma K} n^{-1}\sum_{k=1}^{n}|r_k|\log k \to 0,\ n\to\infty,$$

by the first part of (3.12)'. Since we assume that $r_n \to 0$, there is an integer $N$ such that $\gamma|r_k| < (1-\eta)/2$ for $k \ge N$. Hence

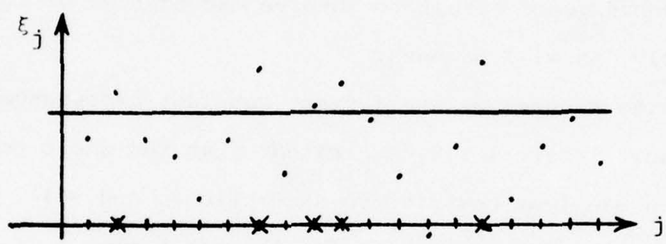
$$n^{-1}\sum_{\substack{k\in\theta_n(K)\\ k\ge N}}|r_k|\log k\, e^{\gamma|r_k|\log k} \le n^{-1}\nu_n(K)\log n\, n^{(1-\eta)/2},$$

which tends to zero as $n\to\infty$, by the second part of (3.12)'. Since $N$ is fixed, $n^{-1}\sum_{k=1}^{N}|r_k|\log k\exp(\gamma|r_k|\log k)\to 0$, and it follows that also the second term of the right hand side of (3.13) tends to zero, and thus that (3.11) is satisfied. □

# CHAPTER 4

## CONVERGENCE OF THE POINT PROCESS OF EXCEEDANCES, AND THE DISTRIBUTION OF r-th LARGEST MAXIMA

In this chapter we return to the general situation and notation of Chapter 2. If $u$ is a given "level" we say that the (stationary) sequence $\{\xi_n\}$ has an *exceedance* of $u$ at $j$, if $\xi_j > u$. We may regard such $j$ as "instants of time", and the exceedances therefore as events ocurring randomly in time, i.e. as a *point process*. (The formal properties of point processes which we shall need are contained in the appendix.)




Point process of exceedances

We shall be concerned with such exceedances for (typically) increasing levels and will define such a point process, $N_n$, say, for each of a sequence $\{u_n\}$ of levels. Since the $u_n$ will be typically high for large n, the exceedances will tend to be rarer and we shall find it convenient to normalize the "time" axis to keep the expected number of exceedances approximately constant. For our purposes the simple scale change by the factor $n$ will suffice. Specifically we define for each $n$ a process $\eta_n(t)$ at the points $t = j/n$, $j = 1, 2, \ldots$ by $\eta_n(j/n) = \xi_j$. Then $\eta_n$ has an exceedance of $u_n$ at $j/n$ whenever $\{\xi_n\}$ has an exceedance at $j$. Hence, while exceedances of $u_n$ may be lost as $u_n$ increases, this will be balanced by the fact that the points $j/n$ become more dense. Indeed the expected number of exceedances by $\eta_n$ in the interval $[0,1]$ is clearly $n\,P\{\xi_1 > u_n\}$ which tends to a finite value $\tau$, if $u_n$ is chosen by (2.9).

Our first task will be to show that (under $D(u_n)$, $D'(u_n)$ conditions) the exceedances of $u_n$ by $\eta_n$ become Poisson in character as $n$ increases, (actually in the full sense of distributional convergence for point processes described in the appendix). In particular this will mean that the number, $N_n(B)$, say, of exceedances of $u_n$ by $\eta_n$ in the (Borel) set $B$, will have an asymptotic Poisson distribution. From this we may simply obtain the asymptotic distribution of the $r$ - th largest among $\xi_1$, ..., $\xi_n$, and thus generalize Theorems 1.13 and 1.14.

It will also be of interest to generalize Theorems 1.15 and 1.16, giving joint distributions of $r$ - th largest values. This will require an extension of our convergence result to involve exceedances of several levels simultaneously, as will be seen.

In the following theorem we shall first consider exceedances of $u_n$ by $\eta_n$ on the unit interval $(0, 1]$ rather than the whole positive axis, since we can then use less restrictive assumptions, and still obtain the corollaries concerning the distributions of $r$ - th maxima.

<u>THEOREM 4.1</u> (i) Let $\tau > 0$ be fixed and suppose that $D(u_n)$, $D'(u_n)$ hold for the stationary sequence $\{\xi_n\}$ with $u_n = u_n(\tau)$ satisfying (2.9). Let $\eta_n(j/n) = \xi_j$, $j = 1, 2, \ldots, n = 1, 2, \ldots$, and let $N_n$ be the point process on the unit interval $(0, 1]$ consisting of the exceedances of $u_n$ by $\eta_n$ in that interval, (i.e. the points $j/n$, $1 \leq j \leq n$, for which $\eta_n(j/n) = \xi_j > u_n$). Then $N_n$ converges in distribution to a Poisson process $N$ on $(0, 1]$ with parameter $\tau$, as $n \to \infty$.

(ii) Suppose that, for each $\tau > 0$, there exists a sequence $\{u_n(\tau)\}$ satisfying (2.9), and that $D(u_n(\tau))$, $D'(u_n(\tau))$ hold for all $\tau > 0$. Then for any fixed $\tau$, the result of (i) holds for the entire positive axis in place of the unit interval, i.e. the point process $N_n$ of exceedances of $u_n(\tau)$ by $\eta_n$, converges to a Poisson Process $N$ on $(0,\infty)$ with parameter $\tau$.

<u>PROOF</u> By Theorem A.1 it is sufficient to show that

(a) $E(N_n((a, b])) \to E(N((a, b])) = \tau(b - a)$ as $n \to \infty$ for all $0 < a < b \leq 1$, and

(b) $P\{N_n(B)=0\} \to P\{N(B)=0\} = e^{-\tau m(B)}$ (m being Lebesgue measure) for all B of the form $\bigcup_1^r (a_i, b_i]$, $0 < a_1 < b_1 < a_2 < \ldots < a_r < b_r \le 1$.

Here (a) is immediate since

$$E(N_n((a, b])) = ([nb]-[na])(1-F(u_n)) \sim n(b-a)\tau/n = \tau(b-a).$$

To show (b) we note, using stationarity, that for $0 < a < b \le 1$,

$$P\{N_n((a, b]) = 0\} = P\{M(I_n) \le u_n\},$$

where $I_n = \{1, 2, \ldots, [bn]-[an]\}$. Now $I_n$ contains $\nu_n$ integers where $\nu_n = [bn]-[an] \sim (b-a)n$ as $n \to \infty$. Further, by Lemma 2.8, since $D(u_n(\tau))$, $D'(u_n(\tau))$ hold they also hold with $u_n(\theta\tau)$ defined by (2.16) replacing $u_n(\tau)$, for $0 < \theta < 1$, and this $u_n(\theta\tau)$ satisfies (2.13) Thus Corollary 2.7 (with $\theta = b-a$) gives

(4.1) $$P\{N_n((a, b]) = 0\} \to e^{-\tau(b-a)} \quad \text{as } n \to \infty.$$

Now, let $B = \bigcup_1^r (a_i, b_i]$, where $0 < a_1 < b_1 < a_2 < b_2 < \ldots < a_r < b_r \le 1$. Then, writing $E_j$ for the set of integers $([na_j]+1, [na_j]+2, \ldots, [nb_j])$, it is readily checked that

$$\begin{aligned} P\{N_n(B) = 0\} &= P\Big(\bigcap_{j=1}^r \{M(E_j) \le u_n\}\Big) \\ &= \prod_{j=1}^r P\{N_n((a_j, b_j]) = 0\} \\ &\quad + \Big\{P\Big(\bigcap_{j=1}^r \{M(E_j) \le u_n\}\Big) - \prod_{j=1}^r P\{M(E_j) \le u_n\}\Big\}. \end{aligned}$$

By (4.1), the first term converges, as $n \to \infty$, to $\prod_{j=1}^r e^{-\tau(b_j - a_j)} = e^{-\tau m(B)}$, (where m denotes Lebesgue measure). On the other hand, by Lemma 2.1, it is readily seen that the modulus of the remaining difference of terms does not exceed $(r-1)\alpha_{n,n\lambda}$ where $\lambda = \min_{1 \le j \le r-1} (a_{j+1} - b_j)$. But by $D(u_n)$, (cf. (2.3)), $\alpha_{n,n\lambda} \to 0$ as $n \to \infty$ so that (b) follows. Hence (i) of the theorem holds.

The conclusion (ii) follows by exactly the same proof except that we require $D(u_n(\theta\tau))$, $D'(u_n(\theta\tau))$ to hold for all $\theta > 1$ as well as $\theta < 1$.

Note that since (ii) and (iii) of Lemma 2.8 require $\theta < 1$, the assumption has to be made in (ii) of the present theorem that $D(u_n(\tau))$, $D'(u_n(\tau))$ hold for all $\tau > 0$. □

Note that in (ii) of the above theorem, we could have assumed the existence of $u_n(\tau)$ satisfying (2.9) for just one fixed $\tau$ and then defined $u_n(\tau')$ for all other $\tau' = \theta\tau$ by $u_n(\theta\tau) = u_{[n/\theta]}(\tau)$ as in Lemma 2.8. Of course it is possible that this may be less convenient in verifying the $D$ and $D'$ conditions than some other prescription of $u_n(\tau)$ satisfying (2.9) for each $\tau$.

It is of interest to note that the conclusion of (i) applies to any interval of unit length, so that the exceedances in *any* such interval become Poisson in character. But if the assumption of (ii) is not made, it may not be true that the exceedances become Poisson on the *entire* axis (or on an interval of greater than unit length).

COROLLARY 4.2 Under the conditions of (i) of the theorem, if $B \subset (0, 1]$ is any Borel set whose boundary has Lebesgue measure zero, $(m(\partial B) = 0)$, then

$$P\{N_n(B) = r\} \to e^{-\tau m(B)}(\tau m(B))^r/r!, \quad r = 0, 1, 2, \ldots$$

The joint distribution of any finite number of variables $N_n(B_1), \ldots, N_n(B_k)$ corresponding to disjoint $B_i$, (with $m(\partial B_i) = 0$ for each $i$) converges to the product of corresponding Poisson probabilities.

PROOF This follows at once since $(N_n(B_1), \ldots, N_n(B_k)) \xrightarrow{d} (N(B_1), \ldots, N(B_k))$ (as noted in the appendix) when $N_n \xrightarrow{d} N$. □

It should be noted that the above results obviously apply very simply to stationary normal sequences satisfying appropriate covariance conditions (e.g. (3.1)).

Asymptotic distribution of k - th largest values

The following results may now be obtained from Corollary 4.2, generalizing the conclusions of Theorems 1.13 and 1.14.

THEOREM 4.3 Let $M_n^{(k)}$ denote the k - th largest of $\xi_1, \ldots, \xi_n$, ($M_n^{(1)} = M_n$), where k is a fixed integer. Let $\{u_n\}$ be a real sequence and suppose that $D(u_n)$, $D'(u_n)$ hold. If (2.9) holds for some fixed $\tau \geq 0$, then

$$(4.2) \qquad P\{M_n^{(k)} \leq u_n\} \to e^{-\tau} \sum_{s=0}^{k-1} \tau^s/s! \quad \text{as} \quad n \to \infty.$$

Conversely if (4.2) holds for some integer k, so does (2.9), and hence (4.2) holds for all k.

PROOF As previously, we identify the event $\{M_n^{(k)} \leq u_n\}$ with the event that no more than $(k-1)$ of $\xi_1, \ldots, \xi_n$ exceed $u_n$, i.e. with $N_n((0, 1]) \leq k-1$, so that

$$(4.3) \qquad P\{M_n^{(k)} \leq u_n\} = \sum_{s=0}^{k-1} P\{N_n((0, 1]) = s\}.$$

If (2.9) holds the limit on the right of (4.2) follows at once by Corollary 4.2.

Conversely if (4.2) holds but (2.9) does not, there is some $\tau' \neq \tau$, $0 \leq \tau' \leq \infty$, and a sequence $\{n_j\}$ such that $n_j(1 - F(u_{n_j})) \to \tau'$. Now a brief examination of the proof of Theorem 4.1 (cf. also the remark after Theorem 2.6) shows that if (2.9) is not assumed for all n but just for a sequence $\{n_j\}$ then $N_{n_j}$ has a Poisson limit. If $\tau' < \infty$, replacing $\tau$ by $\tau'$ in the argument above, we thus have

$$P\{M_{n_j}^{(k)} \leq u_{n_j}\} \to e^{-\tau'} \sum_{s=0}^{k-1} \tau'^s/s!$$

But this contradicts (4.2) since the function $e^{-x} \sum_{s=0}^{k-1} x^s/s!$ is strictly decreasing in $x \geq 0$ and hence 1 - 1. Thus $\tau' < \infty$ is not possible. But $\tau' = \infty$ cannot hold either since as in (2.11) we would have

$$P\{M_{[n/k]} \leq u_n\} \leq 1 - [n/k](1 - F(u_n)) + S_{n,k},$$

which would be negative for large $n$ by the finiteness of $\limsup_{n\to\infty} S_{n,k}$ implied by $D'(u_n)$, at least for some appropriately chosen large $k$. Hence (2.9) holds as asserted. □

The case $k = 1$ of this theorem is just Theorem 2.6 again. Of course Theorem 2.6 is used in the proof of Theorem 4.1 and hence of Theorem 4.3. The following obvious corollary also holds.

<u>COROLLARY 4.4</u> The theorem holds if the assumption (or conclusion) that $\{u_n\}$ satisfy (2.9), is replaced by either of the assumptions (conclusions)

$$P\{M_n \le u_n\} \to e^{-\tau}, \quad P\{\hat{M}_n \le u_n\} \to e^{-\tau}$$

(where $\hat{M}_n$ as usual denotes the sequence of maxima for the associated independent sequence).

<u>PROOF</u> This follows at once from Theorem 2.6. □

<u>THEOREM 4.5</u> Let $a_n > 0$, $b_n$ be constants for $n = 1, 2, \ldots$, and $G$ a non-degenerate d.f., and suppose that $D(u_n)$, $D'(u_n)$ hold for all $u_n = x/a_n + b_n$, $-\infty < x < \infty$. If

(4.4) $\quad P\{a_n(M_n - b_n) \le x\} \xrightarrow{w} G(x)$,

then for each $k = 1, 2, \cdots$

(4.5) $\quad P\{a_n(M_n^{(k)} - b_n) \le x\} \xrightarrow{w} G(x) \sum_{s=0}^{k-1} (-\log G(x))^s/s!$

where $G(x) > 0$ (zero where $G(x) = 0$).

Conversely if (4.5) holds for some $k$, so does (4.4) and hence (4.5) holds for all $k$. Further, the result remains true if $\hat{M}_n$ replaces $M_n$ in (4.4).

<u>PROOF</u> For $G(x) > 0$ the result follows from the previous corollary. The case $G(x) = 0$ follows by continuity of $G$ (which must be an extreme value type), as usual. □

Independence of maxima in disjoint intervals

As noted already in Chapter 2, it would be natural to extend the theorems there to deal with the joint behaviour of maxima in disjoint intervals. We shall do this here – demonstrating asymptotic independence under an appropriate generalization of the $D(u_n)$ condition, and then use this to obtain a Poisson result for exceedances of several levels considered jointly. This, in turn, will lead to the asymptotic joint distributions of various quantities of interest, such as two or more $M_n^{(k)}$, and their locations, as $n \to \infty$.

As in (1.26) we shall consider $r$ levels $u_n^{(k)}$ satisfying

(4.6) $$1 - F(u_n^{(k)}) = \tau_k/n + o(1/n),$$

where $u_n^{(1)} \geq u_n^{(2)} \geq \ldots \geq u_n^{(r)}$ as in Chapter 1 and consequently $\tau_1 \leq \tau_2 \leq \ldots \leq \tau_r$.

It is intuitively clear that we shall need to extend the $D(u_n)$ conditions to involve the $r$ values $u_n^{(k)}$, and we do so as follows.

CONDITION $D_r(\underline{u}_n)$ will be said to hold for the stationary sequence $\{\xi_j\}$ if, for each choice of $\underline{i} = (i_1, \ldots, i_p)$, $\underline{j} = (j_1, \ldots, j_q)$, $1 \leq i_1 < i_2 < \ldots < i_p < j_1 < j_2 < \ldots < j_q \leq n$, $j_1 - i_p \geq \ell$, we have (using obvious notation)

(4.7) $$|F_{\underline{i},\underline{j}}(\underline{v},\underline{w}) - F_{\underline{i}}(\underline{v})\, F_{\underline{j}}(\underline{w})| \leq \alpha_{n,\ell}$$

where $\underline{v} = (v_1, \ldots, v_p)$, $\underline{w} = (w_1, \ldots, w_q)$, the $v_i$ and $w_j$ each being any choice of the $r$ values $u_n^{(1)}, \ldots, u_n^{(r)}$, and where $\alpha_{n,\ell_n} \to 0$ for some sequence $\ell_n = o(n)$.

Condition $D_r(\underline{u}_n)$ extends $D(u_n)$ in an obvious way and will be convenient even though its full strength will not quite be needed for our purposes. It will not be necessary to define an extended $D'(u_n)$ condition, since we shall simply need to assume that $D'(u_n^{(k)})$ holds separately for each $k = 1, 2, \ldots, r$.

The next result extends Lemma 2.1 (with slight changes of notation).

LEMMA 4.6 With the above notation, if $D_r(\underline{u}_n)$ holds, if $n, s, k$ are fixed integers, and $E_1, \ldots, E_s$ subintervals of $\{1, \ldots, n\}$ such that $E_i$ and $E_j$ are separated by at least $\ell$ when $i \neq j$, then

$$\left|P\left(\bigcap_{j=1}^{s}\{M(E_j) \leq u_{n,j}\}\right) - \prod_{j=1}^{s} P\{M(E_j) \leq u_{n,j}\}\right| \leq (s-1)\alpha_{n,\ell}$$

where for each $j$, $u_{n,j}$ is any one of $u_n^{(1)}, \ldots, u_n^{(r)}$.

PROOF This is proved in exactly the same manner as Lemma 2.1 and the details will therefore not be repeated here. □

In the following discussion we shall consider a fixed number $s$ of disjoint subintervals $J_1, J_2, \ldots, J_s$ of $\{1, \ldots, n\}$ such that $J_k$ $(= J_{k,n})$ has $\nu_{k,n} \sim \theta_k n$ elements, where $\theta_k$ are fixed positive constants with $\Sigma_{k=1}^{s}\theta_k \leq 1$. By slightly strengthening the assumptions, we may also allow $\Sigma_{k=1}^{s}\theta_k > 1$, and let $J_1, J_2, \ldots, J_s$ be arbitrary finite disjoint intervals of positive integers. Note that the intervals $J_k$ do increase in size with $n$, but remain disjoint, and their total number $s$ is fixed.

The following results then hold. In the proofs, details will be omitted where they duplicate arguments given in Chapter 2.

THEOREM 4.7 (i) Let $J_1, J_2, \ldots, J_s$ be disjoint subintervals of $\{1, 2, \ldots, n\}$ as defined above, $J_k$ having $\nu_{n,k} \sim \theta_k n$ members, for fixed positive $\theta_1, \theta_2, \ldots, \theta_s$, $(\Sigma_1^s \theta_k \leq 1)$. Suppose that the stationary sequence $\{\xi_j\}$ satisfies $D_r(\underline{u}_n)$ where the levels $u_n^{(1)} \geq u_n^{(2)} \geq \ldots \geq u_n^{(r)}$ satisfy (4.6). Then

$$(4.8) \qquad P\left(\bigcap_{k=1}^{s}\{M(J_k) \leq u_{n,k}\}\right) - \prod_{k=1}^{s} P\{M(J_k) \leq u_{n,k}\} \to 0 \quad \text{as} \quad n \to \infty$$

for any choice of $u_{n,k}$ from $u_n^{(1)}, \ldots, u_n^{(r)}$ for each $k$.

(ii) Suppose, for fixed $\tau_1, \tau_2, \ldots, \tau_r$, that $u_n(\theta\tau_k)$ satisfy (4.6) for all $\theta > 0$, and that $D_r(\underline{u}_n)$ holds for $\underline{u}_n = (u_n(\theta\tau_1), u_n(\theta\tau_2), \ldots, u_n(\theta\tau_r))$, $\theta > 0$. Then (4.8) holds for $J_1, J_2, \ldots, J_s$ finite

disjoint intervals of positive integers with $\nu_{n,k} \sim \theta_k n$ members, where $\theta_1, \theta_2, \ldots, \theta_s$ are arbitrary but fixed positive constants.

PROOF Let $I_k$ denote the first $\nu_{n,k} - \ell_n$ elements of $J_k$, and $I_k^*$ the remaining $\ell_n$, where $\ell_n$ is chosen as in $D_r(\underline{u}_n)$. (These $I_k$, $I_k^*$ are different from those in Chapter 2, but are so named since they play a similar role.) By familiar calculation we have

$$(4.9) \quad 0 \le P\left(\bigcap_{k=1}^{s} \{M(I_k) \le u_{n,k}\}\right) - P\left(\bigcap_{k=1}^{s} \{M(J_k) \le u_{n,k}\}\right)$$
$$\le \sum_{k=1}^{s} P\{M(I_k^*) \ge u_{n,k}\} \le s\rho_n,$$

where $\rho_n = \max_{1\le k\le s} P\{M(I_k^*) \ge u_{n,k}\}$. Further

$$(4.10) \quad \left|P\left(\bigcap_{k=1}^{s} \{M(I_k) \le u_{n,k}\}\right) - \prod_{k=1}^{s} P\{M(I_k) \le u_{n,k}\}\right| \le (s-1)\alpha_{n,\ell_n}$$

by Lemma 4.6, and

$$(4.11) \quad 0 \le \prod_{k=1}^{s} P\{M(I_k) \le u_{n,k}\} - \prod_{k=1}^{s} P\{M(J_k) \le u_{n,k}\}$$
$$\le \prod_{k=1}^{s} (P\{M(J_k) \le u_{n,k}\} + \rho_n) - \prod_{k=1}^{s} P\{M(J_k) \le u_{n,k}\}$$
$$\le (1+\rho_n)^s - 1 \le 2^s \rho_n$$

since $\prod_{k=1}^{s} (y_k + \rho_n) - \prod_{k=1}^{s} y_k$ is increasing in each $y_k$ when $\rho_n > 0$, and we also have $\rho_n \le 1$. Now

$$\rho_n = \max_{1\le k\le s} P\{M(I_k^*) \ge u_{n,k}\} \le \max_{1\le k\le s} \ell_n(1 - F(u_{n,k}))$$

which tends to zero by (4.6) since $\ell_n = o(n)$. Hence, by (4.9), (4.10), (4.11), the left hand side of (4.8) is dominated in absolute value by $(s+2^s)\rho_n + (s-1)\alpha_{n,\ell_n}$ which tends to zero, completing the proof of part (i) of the theorem.

Part (ii) follows by exactly the same arguments under the extended conditions. □

Note that the proof of this theorem is somewhat simpler than that e.g. in Lemmas 2.2 and 2.3. This occurs because we assume (4.6), whereas the corresponding assumption was not made there. We could dispense with (4.6) here also with a corresponding increase in complexity, but since we assume (4.6) in the sequel, we use it here also.

COROLLARY 4.8 (i) If, in addition to the assumptions of part (i) of the theorem, we suppose that $D'(u_n^{(k)})$ holds for each $k = 1, 2, \ldots, r$, then (for $\Sigma_1^s \theta_k \leq 1$),

$$P\{M(J_k) \leq u_{n,k},\ k = 1, 2, \ldots, s\} \to e^{-\sum_{k=1}^{s} \theta_k \tau_k'}$$

where $\tau_k'$ is that one of $\tau_1, \ldots, \tau_r$ corresponding to $u_{n,k}$, i.e. such that $n(1 - F(u_{n,k})) \to \tau_k'$.

(ii) If in addition to the assumptions of Part (ii) of the theorem, $D'(u_n^{(j)})$ are satisfied with $u_n(\theta_j \tau_j)$ in place of $u_n^{(j)}$, $j = 1, \ldots, s$, then the conclusion holds for all constants $\theta_k > 0$.

PROOF This follows by Corollary 2.7 and Lemma 2.8 which show that

$$P\{M(J_k) \leq u_{n,k}\} \to e^{-\theta_k \tau_k'}. \qquad \square$$

It is easy to check that $D_r(\underline{u}_n)$ holds for normal sequences under the standard covariance conditions and hence Corollary 4.8 may be applied.

THEOREM 4.9 Let $\{\xi_n\}$ be a stationary normal sequence with zero means, unit variances and covariance sequence $\{r_n\}$. Let $u_n^{(1)} \geq u_n^{(2)} \geq \cdots \geq u_n^{(r)}$ be $r$ levels such that $(1 - \Phi(u_n^{(k)})) \sim \tau_k/n$, for constants $\tau_k > 0$. Suppose that $r_n \to 0$ and

$$n \sum_{j=1}^{n} |r_j|\, e^{-(u_n^{(r)})^2/(1 + |r_j|)} \to 0 \quad \text{as } n \to \infty, \tag{4.12}$$

which will hold in particular, by Lemma 3.1, if $r_n \log n \to 0$. Then $D_r(\underline{u}_n)$ holds, as do $D'(u_n^{(k)})$ for $1 \leq k \leq r$, and hence, from Corollary 4.8(ii),

(4.13) $P\{M(J_k) \le u_{n,k}\ ,\ k = 1, 2, \ldots, s\} \to e^{-\sum_{k=1}^{s} \theta_k \tau'_k}$

where $J_k$ are as in Theorem 4.7(ii) and $u_{n,k}, \tau'_k$ as in Corollary 4.8.

PROOF With the notation of (4.7), we may identify the $\xi_j$ of Lemma 3.2 with $\xi_{i_1}, \ldots, \xi_{i_p}, \xi_{j_1}, \ldots, \xi_{j_q}$ here and the $\eta_j$ of that lemma with $\xi'_{i_1}, \ldots, \xi'_{i_p}, \xi'_{j_1}, \ldots, \xi'_{j_q}$, such that $\xi'_{i_1}, \ldots, \xi'_{i_p}$ have the same joint distribution as $\xi_{i_1}, \ldots, \xi_{i_p}$, but are independent of $\xi'_{j_1}, \ldots \xi'_{j_q}$ which in turn have the same joint distribution as $\xi_{j_1}, \ldots, \xi_{j_q}$. Then the lemma gives

$$|F_{\underline{i},\underline{j}}(\underline{v},\underline{w}) - F_{\underline{i}}(\underline{v})\ F_{\underline{j}}(\underline{w})| \le K \sum_{\substack{1\le s\le p\\ 1\le t\le q}} |r_{i_s - j_t}|\ e^{-u_n^2/(1 + |r_{i_s - j_t}|)}$$

where $u_n = \min(v_1, \ldots, v_p, w_1, \ldots, w_q)$ and $v_1, \ldots, v_p, w_1, \ldots, w_q$ are chosen among $u_n^{(1)}, \ldots, u_n^{(r)}$. Replacing $u_n$ by $u_n^{(r)}$ $(\le u_n)$ and using the fact that for each $j$ there are at most $n$ terms containing $r_j$, we obtain

$$|F_{\underline{i},\underline{j}}(\underline{v},\underline{w}) - F_{\underline{i}}(\underline{v})F_{\underline{j}}(\underline{w})| \le Kn \sum_{j=1}^{n} |r_j|\ e^{-(u_n^{(r)})^2/(1+|r_j|)}$$

which tends to zero by (4.12), so that $D_r(\underline{u}_n)$ holds as claimed.

Finally Lemma 3.3 shows that $D'(u_n^{(k)})$ holds for each $k$ and hence the final conclusion follows from Corollary 4.8. □

We may combine Theorem 2.9 with Corollary 4.8 in the following obvious way.

THEOREM 4.10 Let $\{\xi_n\}$ be a stationary sequence, $a_n > 0$, $b_n$, constants, and suppose that

$$P\{a_n(M_n - b_n) \le x\} \xrightarrow{w} G(x) \quad \text{as} \quad n \to \infty$$

for some non-degenerate d.f. $G$. Suppose that $D_r(\underline{u}_n)$, $D'(u_n^{(k)})$ hold for all sequences of the form $u_n^{(k)} = x_k/a_n + b_n$, and let $J_k = J_{n,k}$, $k = 1, 2, \ldots, r$ be disjoint subintervals of $\{1, \ldots, n\}$, $J_k$ containing $\nu_{n,k}$

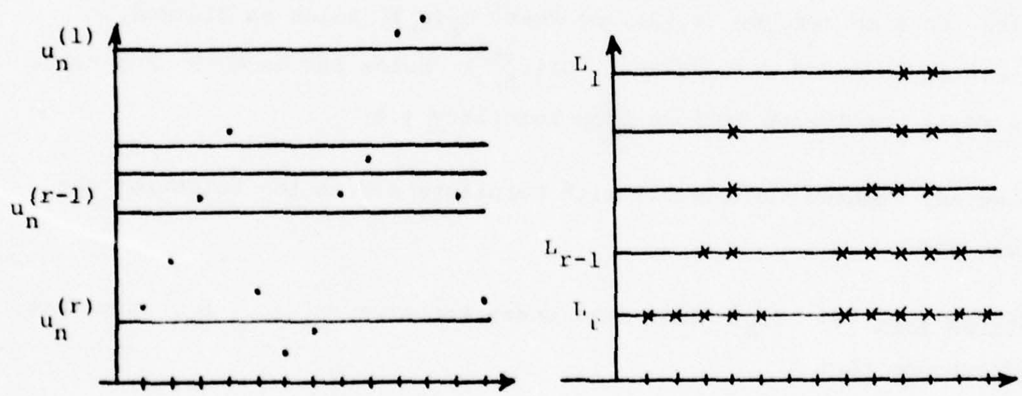
integers where $\nu_{n,k}/n \to \theta_k > 0$, $(\sum_1^r \theta_k \le 1)$. Then the maxima $M(J_k)$ are asymptotically independent. Indeed

$$P\{a_n(M(J_k) - b_n) \le x_k,\ k = 1, 2, \dots, r\} \to \prod_{k=1}^{r} G^{\theta_k}(x_k).$$

PROOF This follows at once from Corollary 4.8(i) and Theorem 2.9 by identifying $\tau'_k$ with $-\log G(x_k)$ and $u_{n,k} = u_n^{(k)} = x_k/a_n + b_n$. □

## Exceedances of multiple levels

It is natural to consider exceedances of the levels $u_n^{(1)}, \dots, u_n^{(k)}$ by $\eta_n$, $(\eta_n(j/n) = \xi_j$ as before), as a vector of point processes. While this may be achieved abstractly, we shall here, as an obvious aid to intuition, represent them as occurring along fixed horizontal lines $L_1, \dots, L_r$ in the plane – exceedances of $u_n^{(k)}$ being represented as points on $L_k$. This will show the structure imposed by the fact that an exceedance of $u_n^{(k)}$ is automatically an exceedance of $u_n^{(k+1)}, \dots, u_n^{(r)}$, as illustrated in the following diagrams.




(i) Levels and values of $\eta_n(t)$ (ii) Representation in plane (fixed $L_i$)

In the diagrams, (i) shows the levels and values of $\eta_n$ from which one can see the exceeded levels, while (ii) marks the points of exceedance of each level along the fixed lines $L_1, \ldots, L_r$.

To pursue this a little further, the diagram (ii) represents the exceedances of all the levels as points in the plane. That is we may regard them as a *point process in the plane* if we wish. To be sure, all the points lie only on certain horizontal lines, and a point on any $L_k$ has points directly below it on all lower $L_k$, but nevertheless the positions are stochastic, and do define a two dimensional point process, which we denote by $N_n$.

We may apply convergence theory to this sequence $\{N_n\}$ of point processes and obtain the joint distributional results as a consequence. The position of the lines $L_1, \ldots, L_r$ do not matter as long as they are fixed and in the indicated order. From our previous theory each one-dimensional point process, on a given $L_k$, $N_n^{(k)}$, say, will become Poisson under appropriate conditions. The two-dimensional process indicated will not become Poisson in the plane as is intuitively clear, in view of the structure described. However, the exceedances $N_n^{(k)}$ on $L_k$ form successively more "severely thinned" versions of $N_n^{(r)}$ as k decreases. Of course, these are not thinnings formed by *independent* removal of events, except in the limit where the Poisson process $N^{(k)}$ on $L_k$ may be obtained from $N^{(k+1)}$ on $L_{k+1}$ by independent thinning, as will become apparent.

More specifically, we define the point process N in the plane, which will turn out to be the appropriate limiting point process, as follows.

Let $\{\sigma_{1j};\ j = 1, 2, \ldots\}$ be the points of a Poisson process with parameter $\tau_r$ on $L_r$. Let $\beta_j$, $j = 1, 2, \ldots$, be i.i.d. random variables, independent also of the Poisson process on $L_r$, taking values $1, 2, \ldots, r$ with probabilities

$$
\begin{aligned}
P\{\beta_j = s\} &= (\tau_{r-s+1} - \tau_{r-s})/\tau_r \qquad s = 1, 2, \ldots, r-1, \\
&= \tau_1/\tau_r \qquad s = r,
\end{aligned}
$$

i.e. $P\{\beta_j \geq s\} = \tau_{r-s+1}/\tau_r$ for $s = 1, 2, \ldots, r$.

For each $j$, place points $\sigma_{2j}, \sigma_{3j}, \ldots, \sigma_{\beta_j j}$ on the $\beta_j-1$ lines $L_{r-1}, L_{r-2}, \ldots, L_{r-\beta_j+1}$, vertically above $\sigma_{1j}$, to complete the point process $N$. Clearly the probability that a point appears on $L_{r-1}$ above $\sigma_{1j}$ is just $P\{\beta_j \geq 2\} = \tau_{r-1}/\tau_r$ and the deletions are independent, so that $N^{(r-1)}$ is obtained as an independent thinning of the Poisson process $N^{(r)}$. Hence $N^{(r-1)}$ is a Poisson process (cf. Appendix) with intensity $\tau_r(\tau_{r-1}/\tau_r) = \tau_{r-1}$, as expected. Similarly $N^{(k)}$ is obtained as an independent thinning of $N^{(k+1)}$ with deletion probability $1 - \tau_k/\tau_{k+1}$, all $N^{(k)}$ being Poisson.

We may now give the main result.

THEOREM 4.11 (i) Suppose that $D_r(\underline{u}_n)$ holds, and that $D'(u_n^{(k)})$ holds for $1 \leq k \leq r$, where the $u_n^{(k)}$ satisfy (4.6). Then the point process $N_n$ of exceedances of the levels $u_n^{(1)}, \ldots, u_n^{(r)}$ (represented as above on the lines $L_1, \ldots, L_r$) converges to the limiting point process $N$, as point processes on $(0, 1] \times R$.

(ii) If the conditions of Corollary 4.8(ii) are satisfied, then $N_n$ converges to $N$, as point processes on the entire right half plane, i.e. on $(0, \infty) \times R$.

PROOF Again by Theorem A.1 it is sufficient to show that

(a) $E(N_n(B)) \to E(N(B))$ for all sets $B$ of the form $(a, b] \times (\alpha, \beta]$, $\alpha < \beta$, $0 < a < b$, where $b \leq 1$ or $b < \infty$, respectively, and

(b) $P\{N_n(B) = 0\} \to P\{N(B) = 0\}$ for all sets $B$ which are finite unions of disjoint sets of this form.

Again (a) follows readily. If $B = (a, b] \times (\alpha, \beta]$ intersects any of the lines, let these be $L_s, L_{s+1}, \ldots, L_t$, $(1 \leq s \leq t \leq r)$. Then $N_n(B) = \sum_{k=s}^{t} N_n^{(k)}((a, b])$, $N(B) = \sum_{k=s}^{t} N^{(k)}((a, b])$ so that

$$E(N_n(B)) = ([nb] - [na]) \sum_{k=s}^{t} (1 - F(u_n^{(k)}))$$

$$\sim n(b-a)\sum_{k=s}^{t}(\tau_k/n + o(1/n)) \to (b-a)\sum_{k=s}^{t}\tau_k$$

which is clearly just $E(N(B))$.

To show (b) we must prove that $P\{N_n(B) = 0\} \to P\{N(B) = 0\}$ for sets $B$ of the form $B = \bigcup_1^m C_k$ with $C_k = (a_k, b_k] \times (\alpha_k, \beta_k]$. By considering intersections and differences of the intervals $(a_k, b_k]$, we may change these so that, for $k \neq \ell$, $(a_k, b_k]$ and $(a_\ell, b_\ell]$ are either identical or disjoint. If these intervals are identical, $(\alpha_k, \beta_k]$ and $(\alpha_\ell, \beta_\ell]$ must be disjoint, so that $\alpha_\ell > \beta_k$ or $\alpha_k > \beta_\ell$. If, say, the first of these holds, $N_n(C_k) = 0$ implies $N_n(C_\ell) = 0$ (by the "thinning property" of the $N_n$) so that $\{N_n(C_k) = 0\} \cap \{N_n(C_\ell) = 0\} = \{N_n(C_k) = 0\}$, and hence we may write

$$(4.14) \quad \{N_n(B) = 0\} = \bigcap_{k=1}^{s}\{N_n(C_k) = 0\},$$

where $C_k = (a_k, b_k] \times (\alpha_k, \beta_k]$, and the intervals $(a_k, b_k]$ are disjoint. But if the lowest $L_j$ intersecting $C_k$ is $L_{m_k}$, say, clearly

$$(4.15) \quad \{N_n(C_k) = 0\} = \{N_n^{(m_k)}((a_k, b_k]) = 0\}.$$

But this is just the event $\{M([a_k n], [b_k n]) \le u_{m_k}\}$, so that Corollary 4.8, part (i) and (ii), respectively, gives

$$P\{N_n(B) = 0\} \to e^{-\sum_{k=1}^{s}(b_k - a_k)\tau_{m_k}} = P\{N(B) = 0\},$$

since (4.14) and (4.15) clearly also hold with $N$ instead of $N_n$, and the result follows. □

<u>COROLLARY 4.12</u> Let $\{\xi_n\}$ satisfy the conditions of Theorem 4.11(i) or (ii), and let $B_1, \ldots, B_s$ be Borel subsets of the unit interval, or the positive real line, respectively, whose boundaries have zero Lebesgue measure. Then for integers $m_j^{(k)}$,

$$(4.16) \quad P\{N_n^{(k)}(B_j) = m_j^{(k)},\ j=1,\ldots,s,\ k=1,\ldots,r\}$$
$$\to P\{N^{(k)}(B_j) = m_j^{(k)},\ j=1,\ldots,s,\ k=1,\ldots,r\}.$$

PROOF Let $B_{jk}$ be a rectangle in the plane with base $B_j$ and such that $L_k$ intersects its interior, but is disjoint from all other $L_j$. Then the left hand side of (4.16) may be written as

$$P\{N_n(B_{jk}) = m_j^{(k)},\ j=1,\ldots,s,\ k=1,\ldots,r\},$$

which by the appendix converges to the same quantity with $N$ instead of $N_n$, i.e. to the right hand side of (4.16). □

Joint asymptotic distribution of the largest maxima

We may apply the above results to obtain asymptotic joint distributions for a finite number of the k - th largest maxima $M_n^{(k)}$, together with their locations if we wish. Such results may be obtained by considering appropriate continuous functionals of the sequence $N_n$, but here we take a more elementary approach, in giving examples of typical results. First we generalize Theorems 1.15 and 1.16.

THEOREM 4.13 Let the levels $u_n^{(k)}$, $1 \leq k \leq r$, satisfy (4.6) with $u_n^{(1)} \geq u_n^{(2)} \geq \ldots \geq u_n^{(r)}$, and suppose that the stationary sequence $\{\xi_n\}$ satisfies $D_r(\underline{u}_n)$, and $D'(u_n^{(k)})$ for $1 \leq k \leq r$. Let $S_n^{(k)}$ denote the number of exceedances of $u_n^{(k)}$ by $\xi_1, \ldots, \xi_n$. Then, for $k_1 \geq 0, \ldots, k_r \geq 0$,

$$(4.17)\quad P\{S_n^{(1)} = k_1,\ S_n^{(2)} = k_1+k_2,\ \ldots,\ S_n^{(r)} = k_1+\ldots+k_r\}$$

$$\to \frac{\tau_1^{k_1}}{k_1!}\,\frac{(\tau_2-\tau_1)^{k_2}}{k_2!}\cdots\frac{(\tau_r-\tau_{r-1})^{k_r}}{k_r!}\,e^{-\tau_r},$$

as $n \to \infty$.

PROOF With the previous notation $S_n^{(j)} = N_n^{(j)}((0,1])$ and so by Corollary 4.12 the left hand side of (4.17) converges to

$$(4.18)\quad P\{S^{(1)} = k_1,\ S^{(2)} = k_1+k_2,\ \ldots,\ S^{(r)} = k_1+\ldots+k_r\},$$

where $S^{(j)} = N^j((0,1])$. But this is the probability that precisely $k_1+k_2+\ldots+k_r$ events occur in the unit interval for the Poisson process

on the line $L_r$ and that $k_1$ of the corresponding $\beta$'s take the value $r$, $k_2$ take the value $r-1$, and so on. But the independence properties of the $\beta$'s show that, conditional on a given total number $k_1+k_2+\ldots+k_r$, the numbers taking the values $r, r-1, \ldots, 1$ have a multinomial distribution based on the respective probabilities $\tau_1/\tau_r$, $(\tau_2-\tau_1)/\tau_r$, $\ldots$, $(\tau_r-\tau_{r-1})/\tau_r$. Hence (4.18) is

$$\frac{(k_1+k_2\ldots+k_r)!}{k_1!k_2!\ldots k_r!}\left(\frac{\tau_1}{\tau_r}\right)^{k_1}\left(\frac{\tau_2-\tau_1}{\tau_r}\right)^{k_2}\ldots\left(\frac{\tau_r-\tau_{r-1}}{\tau_r}\right)^{k_r}$$
$$\times P\{N^{(r)}((0,1]) = k_1+k_2\ldots+k_r\}$$

which gives (4.17) since

$$P\{N^{(r)}((0,1]) = k_1+k_2\ldots+k_r\} = e^{-\tau_r}\tau_r^{k_1+k_2\ldots+k_r}/(k_1+k_2+\ldots+k_r)! \quad \square$$

Of course this agrees with the result of Theorem 1.15. The next result (which generalizes Theorem 1.16) again is given to exemplify the applicability of the Poisson theory.

<u>THEOREM 4.14</u> Suppose that

(4.19) $\quad P\{a_n(M_n^{(1)}-b_n) \le x\} \to G(x)$

for some non-degenerate d.f. $G$ and that $D_2(\underline{u}_n)$, $D'(u_n^{(k)})$ hold whenever $u_n^{(k)} = x_k/a_n+b_n$, $k = 1, 2$. Then the conclusion of Theorem 1.16 holds, i.e. for $x_1 > x_2$,

(4.20) $\quad P\{a_n(M_n^{(1)}-b_n) \le x_1,\ a_n(M_n^{(2)}-b_n) \le x_2\}$

$$\to G(x_2)(\log G(x_1) - \log G(x_2) + 1)$$

when $G(x_2) > 0$ (zero when $G(x_2) = 0$).

<u>PROOF</u> If $u_n^{(k)} = x_k/a_n+b_n$, by (4.19) and the assumptions of the theorem it follows from Theorem 2.6 that $1-F(u_n^{(k)}) \sim \tau_k/n$ where $\tau_k = -\log G(x_k)$, if $G(x_k) > 0$. But clearly

$$P\{M_n^{(1)} \le u_n^{(1)},\ M_n^{(2)} \le u_n^{(2)}\}$$
$$= P\{S_n^{(1)} = 0,\ S_n^{(2)} = 0\} + P\{S_n^{(1)} = 0,\ S_n^{(2)} = 1\},$$

where $S_n^{(i)}$ is the number of exceedances of $u_n^{(i)}$ by $\xi_1, \ldots, \xi_n$. By Theorem 4.13 we see that the limit of the above probabilities is

$$e^{-\tau_2} + (\tau_2 - \tau_1)e^{-\tau_2} = e^{-\tau_2}(\tau_2 - \tau_1 + 1),$$

which is the desired result. □

As a final example, we obtain the limiting joint distribution of the second maximum and its location, (taking the leftmost if two values are equal).

<u>THEOREM 4.15</u> Suppose that (4.19) holds and that $D_4(\underline{u}_n)$, $D'(u_n^{(k)})$ hold for all $u_n^{(k)} = x_k/a_n + b_n$, $k = 1, 2, 3, 4$. Then if $L_n^{(2)}$, $M_n^{(2)}$ are the location and height of the second largest of $\xi_1, \ldots, \xi_n$, respectively,

(4.21) $$P\{\tfrac{1}{n} L_n^{(2)} \le t,\ a_n(M_n^{(2)} - b_n) \le x\} \to t\, G(x)(1 - \log G(x)),$$

$x$ real, $0 < t < 1$. That is, the location and height are asymptotically independent, the location being asymptotically uniform.

<u>PROOF</u> As in the previous theorem, we see that (4.6) holds with $\tau_k = -\log G(x_k)$. Write $I$, $J$ for the intervals $\{1, 2, \ldots, [nt]\}$, $\{[nt] + 1, \ldots, n\}$, respectively, $M^{(1)}(I)$, $M^{(2)}(I)$, $M^{(1)}(J)$, $M^{(2)}(J)$ for the maxima and second largest $\xi_j$ in the intervals $I$, $J$, and let $H_n(x_1, x_2, x_3, x_4)$ be the joint d.f. of the normalized r.v.'s

$$X_n^{(1)} = a_n(M_n^{(1)}(I) - b_n),\ X_n^{(2)} = a_n(M_n^{(2)}(I) - b_n),$$
$$Y_n^{(1)} = a_n(M_n^{(1)}(J) - b_n),\ Y_n^{(2)} = a_n(M_n^{(2)}(J) - b_n).$$

That is with $x_1 > x_2$ and $x_3 > x_4$

$$H_n(x_1, x_2, x_3, x_4) = P\{M_n^{(1)}(I) \le u_n^{(1)},\ M_n^{(2)}(I) \le u_n^{(2)},$$
$$M_n^{(1)}(J) \le u_n^{(3)},\ M_n^{(2)}(J) \le u_n^{(4)}\},$$

where $u_n^{(k)} = x_k/a_n + b_n$ as above. Alternatively we see that

$$H_n(x_1, x_2, x_3, x_4) = P\{N_n^{(1)}(I') = 0,\ N_n^{(2)}(I') \le 1,$$
$$N_n^{(3)}(J') = 0,\ N_n^{(4)}(J') \le 1\},$$

where $I' = (0, t]$ and $J' = (t, 1]$, so that an obvious application of Corollary 4.12 with $B_1 = I'$, $B_2 = J'$ gives

$$(4.22)\quad \lim_{n\to\infty} H_n(x_1, x_2, x_3, x_4) = P\{N^{(1)}(I') = 0,\ N^{(2)}(I') \le 1\}$$

$$\times P\{N^{(3)}(J') = 0,\ N^{(4)}(J') \le 1\}$$

$$= e^{-t\tau_2}(t(\tau_2-\tau_1)+1)e^{-(1-t)\tau_4}((1-t)(\tau_4-\tau_3)+1)$$

$$= H_t(x_1,x_2)H_{(1-t)}(x_3,x_4) = H(x_1,x_2,x_3,x_4),$$

say, where

$$H_t(x_1, x_2) = G^t(x_2)(\log G^t(x_1) - \log G^t(x_2) + 1),$$

for $x_1 > x_2$. Now clearly

$$(4.23)\quad P\{\tfrac{1}{n} L_n^{(2)} \le t,\ a_n(M_n^{(2)} - b_n) \le x_2\}$$

$$= P\{M_n^{(2)}(I) \le u_n^{(2)},\ M_n^{(2)}(I) \ge M_n^{(1)}(J)\}$$

$$+ P\{M_n^{(1)}(I) \le u_n^{(2)},\ M_n^{(1)}(J) > M_n^{(1)}(I) \ge M_n^{(2)}(J)\}$$

$$= P\{X_n^{(2)} \le x_2,\ X_n^{(2)} \ge Y_n^{(1)}\} + P\{X_n^{(1)} \le x_2,\ Y_n^{(1)} > X_n^{(1)} \ge Y_n^{(2)}\}.$$

But, by the above calculation $(X_n^{(1)}, X_n^{(2)}, Y_n^{(1)}, Y_n^{(2)})$ converges in distribution to $(X_1, X_2, Y_1, Y_2)$, whose joint d.f. is $H$, which is clearly absolutely continuous since $G$ is (being an extreme value d.f.). Hence, since the boundaries of sets in $R^4$ such as $\{(w_1, w_2, w_3, w_4);\ w_2 \le x_2,\ w_2 > w_3\}$ clearly have zero Lebesgue measure, it follows that the sum of probabilities in (4.23) converges to

$$(4.24)\quad P\{X_2 \le x_2,\ X_2 \ge Y_1\} + P\{X_1 \le x_2,\ Y_1 > X_1 \ge Y_2\},$$

and therefore the left hand side of (4.21) converges to (4.24). This may be evaluated using the joint distribution $H$ of $X_1, X_2, Y_1, Y_2$ given by (4.22). However, it is simpler to note that we would obtain the same result (4.24) if the sequence were i.i.d. But (4.21) is simply evaluated for an i.i.d. sequence by noting the independence of $L_n^{(2)}$ and $M_n^{(2)}$ and the fact that $L_n^{(2)}$ is then uniform, giving the limit stated in (4.21), as is readily shown. □

Complete Poisson convergence

In the previous point process convergence results, we obtained a limiting point process in the plane, formed from the exceedances of a fixed number $r$ of increasingly high levels. The limiting process was not Poisson in the plane, though composed of $r$ successively more severely thinned Poisson processes on $r$ lines.

On the other hand we may regard the sample sequence $\{\xi_n\}$ itself - after suitable transformations of both coordinates - as a point process in the plane and, by somewhat strengthening the assumptions, show that this converges to a Poisson process in the plane. This procedure has been used for independent r.v.'s by e.g. Pickands (1971) and Resnick (1975) and more recently for stationary sequences by R.J. Adler (1978), who used the linear normalization of process values provided by the constants $a_n$, $b_n$ appearing in the asymptotic distribution of $M_n$. Here we shall consider a slightly more general case.

Specifically, with the standard notation, suppose that $u_n(\tau)$ is defined for $n = 1, 2, \ldots, \tau > 0$ to be strictly decreasing in $\tau$ for each $n$, and satisfying (1.17), viz.,

$$(4.25) \quad 1 - F(u_n(\tau)) = \tau/n + o(1/n).$$

Here we will use $N_n$ to denote the point process in the plane consisting of the points $(j/n, u_n^{-1}(\xi_j))$, $j = 1, 2, \ldots$, where $u_n^{-1}$ denotes the inverse function of $u_n(\tau)$, (defined on the range of the r.v.'s $\xi_j$).

<u>THEOREM 4.16</u> Suppose $u_n(\tau)$ are defined as above satisfying (4.25) and that $D'(u_n)$ holds for each $u_n = u_n(\tau)$, and that $D_r(\underline{u}_n)$ holds for each $r = 1, 2, \ldots, \underline{u}_n = (u_n(\tau_1), \ldots, u_n(\tau_r))$ for each choice of $\tau_1, \ldots, \tau_r$. Then the point processes $N_n$, consisting of the points $(j/n, u_n^{-1}(\xi_j))$, converge to a Poisson process $N$ on $(0, \infty) \times (0, \infty)$, having Lebesgue measure as its intensity.

<u>PROOF</u> This follows relatively simply by using the previous r- level ex-

ceedance theory. Here it will be convenient to use rectangles whose vertical intervals are closed at the bottom rather than the top, so that we need to show

(a) $E(N_n(B)) \to E(N(B))$ for all sets $B$ of the form $(a, b] \times [\alpha, \beta)$, $0 < a < b$, $0 < \alpha < \beta$, and

(b) $P\{N_n(B) = 0\} \to P\{N(B) = 0\}$ for sets $B$ which are finite unions of sets of this form.

Here, (a) follows simply since if $B = (a, b] \times [\alpha, \beta)$,

$$\begin{aligned} E(N_n(B)) &= ([nb] - [na])\, P\{\alpha \le u_n^{-1}(\xi_1) < \beta\} \\ &\sim n(b-a)\, P\{u_n(\beta) < \xi_1 \le u_n(\alpha)\} \\ &= n(b-a)\, \{F(u_n(\alpha)) - F(u_n(\beta))\} \\ &\sim n(b-a)(\beta-\alpha)/n, \end{aligned}$$

while $E(N(B)) = (b-a)(\beta-\alpha)$.

To show (b) we note that any finite disjoint union of such rectangles may be written in the form $\bigcup_j (E_j \times F_j)$, where $E_j = (a_j, b_j]$ are disjoint and $F_j$ is a finite disjoint union $\bigcup_k [\alpha_{j,k}, \beta_{j,k})$, (cf. the proof of Theorem 4.11). Suppose first that there is just one set $E_j$, i.e. $B = \bigcup_{k=1}^{m} E \times F_k$, say, where we write $F_k = [\tau_{2k-1}, \tau_{2k})$, $k = 1, \ldots, m$, and where we may clearly take $\tau_1 < \tau_2 < \ldots < \tau_r$, $(r = 2m)$.

Now $N_n(B) = 0$ means that, for each $k$, there is no $j/n \in E$ for which $u_n^{-1}(\xi_j) \in F_k$, i.e. such that $u_n(\tau_{2k}) < \xi_j \le u_n(\tau_{2k-1})$. But this is equivalent to the statement that for $j/n \in E$, $\xi_j$ exceeds $u_n(\tau_{2k-1})$ as many times as it exceeds $u_n(\tau_{2k})$. That is, writing $N_n^{(k)}(E)$ for the number of exceedances of $u_n(\tau_k)$ by $\xi_j$ for $j/n \in E$,

$$\{N_n(B) = 0\} = \bigcap_{k=1}^{r/2} \{N_n^{(2k-1)}(E) = N_n^{(2k)}(E)\}. \tag{4.26}$$

But $N_n^{(k)}$ is precisely the same as in Theorem 4.11 and its corollary, and their conditions are clearly satisfied, so that by Corollary 4.12(ii) we have

(4.27) $\quad (N_n^{(1)}(E), N_n^{(2)}(E), \ldots, N_n^{(r)}(E)) \xrightarrow{d} (N^{(1)}(E), N^{(2)}(E), \ldots, N^{(r)}(E)),$

where $N^{(1)}, \ldots, N^{(r)}$ are the $r$ successively thinned Poisson processes on $r$ fixed lines as defined prior to Theorem 4.11. But since all the r.v.'s $N_n^{(k)}(E)$, $N^{(k)}(E)$ are integer valued, it is an obvious exercise in distributional convergence in $R^r$ to show from (4.27) that the probability of pairwise equality in (4.26) converges to the same probability with $N^{(k)}$ replacing $N_n^{(k)}$. Thus

$$P\{N_n(B) = 0\} \to P\Big(\bigcap_{k=1}^{r/2} \{N^{(2k-1)}(E) = N^{(2k)}(E)\}\Big).$$

From the discussion prior to Theorem 4.11 we see that the events in braces on the right occur if the $\beta_j$ corresponding to each Poisson event on the line $L_r$ is even, i.e. $\beta_j = 2, 4, 6, \ldots, r$. Since $P\{\beta_j = s\} = (\tau_{r-s+1} - \tau_{r-s})/\tau_r$ if $s \le r-1$, and $\tau_1/\tau_r$ if $s = r$, we have, writing $\gamma = (\tau_{r-1} - \tau_{r-2}) + (\tau_{r-3} - \tau_{r-4}) + \ldots + (\tau_3 - \tau_2) + \tau_1$,

$$P\{N_n(B) = 0\} \to \sum_{j=0}^{\infty} e^{-\tau_r(b-a)} \frac{(\tau_r(b-a))^j}{j!} \left(\frac{\gamma}{\tau_r}\right)^j$$

$$= e^{(b-a)(\gamma - \tau_r)} = e^{-m(B)},$$

where $m$ denotes Lebesgue measure, since

$$(b-a)(\gamma - \tau_r) = \sum_{k=1}^{r/2} (b-a)(\tau_{2k-1} - \tau_{2k}) = -\sum_{k=1}^{r/2} m(E \times F_k) = -m(B).$$

Hence (b) follows when $B = \bigcup_k E \times F_k$. When $B = \bigcup_j (E_j \times F_j)$ the same proof applies – using the full statement of Corollary 4.12, with slightly more notational complexity since more $\tau_k$'s may be needed corresponding to the additional $E_j$'s. □

In the theorem above it is required, of course that $P\{M_n \le u_n(\tau)\} \to e^{-\tau}$. If there is a d.f. $G(x)$ and $\tau$ may be chosen as a function $\tau(x)$ such that $P\{M_n \le u_n(\tau(x))\} \to G(x)$, then we would have $\tau(x) = -\log G(x)$ and $P\{M_n \le v_n(x)\} \to G(x)$, with $v_n(x) = u_n(\tau(x))$. In such a case it would be natural to consider the point process formed from points $(j/n, v_n^{-1}(\xi_j))$

instead of $(j/n, u_n^{-1}(\xi_j))$. In particular when a linear normalization leads to an asymptotic distribution, i.e. $P\{a_n(M_n - b_n) \le x\} \to G(x)$ we have $v_n(x) = x/a_n + b_n$ and it is natural to consider the point process $N_n'$ consisting of points $(j/n, a_n(\xi_j - b_n))$. This is the case considered in Adler (1978) where it is shown that a (non-homogeneous) Poisson limit holds. Here we obtain this result as a corollary of Theorem 4.16.

THEOREM 4.17 Suppose that (4.4) holds, i.e. $P\{a_n(M_n - b_n) \le x\} \to G(x)$ for some non-degenerate d.f. $G$. Suppose that $D'(u_n)$ holds for all sequences $u_n = x/a_n + b_n$, and that $D_r(\underline{u}_n)$ holds for all $r = 1, 2, \ldots,$ and all sequences $u_n^{(k)} = x_k/a_n + b_n$, $1 \le k \le r$, for arbitrary choices of the $x_k$. Then if $N_n'$ denotes the point process in the plane with points at $(j/n, a_n(\xi_j - b_n))$ we have $N_n' \xrightarrow{d} N'$, where $N'$ is a Poisson process whose intensity measure is the product of Lebesgue measure and that defined by the increasing function $\log G(y)$.

PROOF By Theorem 2.6, the conditions of Theorem 4.16 hold, and hence $N_n \xrightarrow{d} N$, with the notation of Theorem 4.16. But if $N_n$ has an atom at $(s, t)$, $N_n'$ has an atom at $(s, \tau^{-1}(t))$ where $\tau(x) = -\log G(x)$. Hence by Theorem A. $N_n' \xrightarrow{d} N'$ where $N'$ is obtained from the Poisson process $N$ by replacing atoms at points $(s, t)$ by atoms at points $(s, \tau^{-1}(t))$, and by Theorem A.2 , this is also a Poisson process, with intensity measure $\lambda'$ defined by

$$\lambda'((a, b] \times (\alpha, \beta]) = (b - a)(\tau(\alpha) - \tau(\beta)) = (b - a)(\log G(\beta) - \log G(\alpha))$$

from which the result follows. (Note that since $G$ is an extreme value distribution, $\tau$ is continuous and strictly decreasing where $G$ is non-zero.) $\square$

Finally we note that all the results which follow from the multi-level result (Theorem 4.11) may be obtained form the last two theorems – however the D-assumptions made are correspondingly more stringent.

## CHAPTER 5

## NORMAL SEQUENCES UNDER STRONG DEPENDENCE*

In Chapter 3 and 4 it was seen that if the dependence in a stationary normal sequence is not too strong, then the extreme values of the dependent sequence have the same asymptotic distribution as the extreme values of independent normal variables, and it was shown that the exceedances of the level $u_n = x/a_n + b_n$ converge in distribution to a Poisson process with intensity $e^{-x}$. Thus, in particular the distribution of $a_n(M_n - b_n)$ tends to a double exponential distribution and the numbers of exceedances in disjoint intervals are asymptotically independent. The crucial conditions needed for these results concern the behaviour of $r_n \log n$; they hold if $r_n \log n \to 0$ or, in somewhat more general circumstances, if $r_n \log n$ is not too large too often. Recent results by Mittal and Ylvisaker (1975), which will be given below, show that these conditions are almost the best possible ones. Their results are that if $r_n \log n \to \gamma > 0$ then $a_n(M_n - b_n)$ does not tend in distribution to $\exp(-e^{-x})$ but to a convolution of $\exp(-e^{-x})$ and a normal distribution function and further that if $r_n \log n \to \infty$ in a sufficiently smooth manner (but $r_n \to 0$ still) then a different normalization is needed and the limiting distribution is normal.

In this chapter we will use ideas from the paper by Mittal and Ylvisaker to show that if $r_n \log n \to \gamma > 0$ then the point process of exceedances of the level $u_n$ converges weakly to a Cox process (i.e. a mixture of Poisson processes with different intensities). The slow decay of the correlations not only changes the limiting distribution of extremes, but also destroys the asymptotic independence between extreme values in disjoint intervals. The reason for this is explained in an instructive way in the proof of Theorem 5.2 below, where the limiting distribution of the exceedances of $u_n$ is obtained as the limiting distribution of the exceedances

* This chapter contains more special material and can be omitted at first reading.

by an independent normal sequence of a random level $(x + \gamma - \sqrt{2\gamma}\zeta)/a_n + b_n$, where $\zeta$ is a standard normal variable representing "the common part" of the first $n$ dependent variables.

The main tool for the proof will, as in Chapter 3, be Lemma 3.2 which relates the distributions of the maxima of two normal sequences with different correlations. Now it is of course no longer sufficient to compare with an independent sequence. Instead it will be convenient to compare the distribution of $M_n$ with that of the maximum $M_n(\delta)$ of $n$ standard normal variables which have constant covariance $\delta \geq 0$ between any two variables. The usefulness of this comparison stems from the fact that if $\zeta, \zeta_1, \zeta_2, \ldots$ are independent standard normal variables then $(1-\delta)^{1/2}\zeta_1 + \delta^{1/2}\zeta, \ldots, (1-\delta)^{1/2}\zeta_n + \delta^{1/2}\zeta$ have constant covariance $\delta$ between any two, and thus $M_n(\delta)$ has the same distribution as $(1-\delta)^{1/2} M_n(0) + \delta^{1/2}\zeta$. That the comparison is possible follows from the first lemma.

LEMMA 5.1 Let $b > 0$ and $\gamma \geq 0$ be constants, put $\rho_n = \gamma/\log n$ and suppose that

(5.1) $\quad r_n \log n \to \gamma \quad \text{as} \quad n \to \infty.$

Then

(5.2) $$nb \sum_{k=1}^{[nb]} |r_k - \rho_n| e^{-u_n^2/(1+w_k)} \to 0 \quad \text{as} \quad n \to \infty,$$

where $w_k = \max(\rho_n, |r_k|)$.

PROOF Put $\delta(k) = \sup_{k<m\leq[bn]} w_m$ and let $p = [n^\alpha]$, where $0 < \alpha < \frac{1-\delta(0)}{1+\delta(0)}$. As in the proof of Lemma 3.1 the contribution from the sum up to $p$ is tends to zero, so we only have to prove that the remaining part of the sum also tends to zero. Now

(5.3) $$n \sum_{k=p+1}^{[nb]} |r_k - \rho_n| e^{-u_n^2/(1+w_k)} \leq n\, e^{-u_n^2/(1+\delta(p))} \sum_{k=p+1}^{[nb]} |r_k - \rho_n|$$

$$= \frac{n^2}{\log n} e^{-u_n^2/(1+\delta(p))} \frac{\log n}{n} \sum_{k=p+1}^{[nb]} |r_k - \rho_n|.$$

Since $r_n \log n \to \gamma$ there is a constant $C$ such that $r_n \log n \le C,\ n \ge 1$. Hence also $\delta(p) \log p \le C$ so by (3.4) we have (again letting $K$ be a constant whose value may change from line to line)

$$(5.4)\qquad \frac{n^2}{\log n} e^{-u_n^2/(1+\delta(p))} \le \frac{n^2}{\log n} e^{-u_n^2/\left(1+\frac{C}{\log n^\alpha}\right)}$$

$$\sim K \frac{n^2}{\log n}\left(\frac{u_n}{n}\right)^{\frac{2}{1+C/\log n^\alpha}}$$

$$\le K n^{\frac{2C/\log n^\alpha}{1+C/\log n^\alpha}} = O(1)$$

as $n \to \infty$. Moreover, adding and subtracting $\rho_n \log n/\log k = \gamma/\log k$ and using the fact that, $\log k \ge \log n^\alpha$ for $k > p$, gives

$$(5.5)\qquad \frac{\log n}{n}\sum_{k=p+1}^{[nb]} |r_k - \rho_n| \le \frac{1}{\alpha n}\sum_{k=p+1}^{[nb]} |r_k \log k - \gamma| + \gamma\cdot\frac{1}{n}\sum_{k=p+1}^{[nb]} \left|1 - \frac{\log n}{\log k}\right|.$$

Here the first term to the right tends to zero by (5.1). Furthermore, estimating the second sum by an integral, we obtain

$$\frac{1}{n}\sum_{k=p+1}^{[nb]}\left|1 - \frac{\log n}{\log k}\right| \le \frac{1}{\alpha \log n}\sum_{k=p+1}^{[nb]}\left|\log\frac{k}{n}\right|\frac{1}{n}$$

$$= O\left(\frac{1}{\alpha \log n}\int_0^b |\log x|\,dx\right),$$

and hence the left hand side of (5.5) tends to zero. Since by (5.4), the first factor on the right of (5.3) is bounded, this concludes the proof of (5.2). □

It is possible to weaken the hypothesis of Lemma 5.1 and thus of Theorem 5.2 below in the same way as (3.1) is weakened to (3.11). However, this is quite straightforward and is left to the reader.

Before stating the theorem we recall from Chapter 4 the notation $N_n$ for the point process of exceedances of the level $u_n$ by $\eta_n$, where $\eta_n$ is defined from the stationary sequence $\{\xi_j\}$ by $\eta_n(j/n) = \xi_j$, $j = 1, 2, \ldots;\ n = 1, 2, \ldots$ Further, let $N$ be a Cox process with (stochastic) intensity $\exp(-x - \gamma + \sqrt{2\gamma}\zeta)$, where $\zeta$ is a standard normal random variable, i.e. let $N$ have the distribution determined by

$$(5.6)\quad P\Big(\bigcap_{i=1}^{k}\{N(B_i)=k_i\}\Big)$$
$$= \int_{-\infty}^{\infty}\prod_{i=1}^{k}\Big\{\frac{(|B_i|e^{-x-\gamma+\sqrt{2\gamma}z})^{k_i}}{k_i!}\exp\{-|B_i|e^{-x-\gamma+\sqrt{2\gamma}z}\}\Big\}\phi(z)\,dz$$

for $B_1, \ldots, B_r$ positive disjoint Borel sets (and $|B_i|$ denotes Legesgue measure).

THEOREM 5.2 Suppose that $\{\xi_n\}$ is a stationary normal sequence with covariances $\{r_n\}$ and that the levels $u_n = x/a_n + b_n$, with $a_n = (2\log n)^{1/2}$ and $b_n = a_n + (2a_n)^{-1}(\log\log n + \log 4\pi)$. If (5.1) holds then the point process $N_n$ of time-normalized exceedances of the level $u_n$ converges in distribution to $N$, where $N$ is the Cox process defined by (5.6).

PROOF Again we have to verify (a) and (b) of Theorem A.1. As in the proof of Theorem 4.1, $E(N_n((a, b]))$ $(b-a)e^{-x}$, so since

$$E(N((a, b])) = E((b-a)e^{-x-\gamma+\sqrt{2\gamma}\zeta}) = (b-a)e^{-x-\gamma}e^{(\sqrt{2\gamma})^2/2}$$
$$= (b-a)e^{-x},$$

the first condition follows immediately.

We use the notation $M_n(a, b) = \max\{\xi_k;\ [an] < k \le [bn]\}$ and write $M_n(a, b; \rho)$ for the maximum of the variables with index $k$, $[an] < k \le [bn]$, in a normal sequence with constant covariance $\rho$ between any two variables. Letting $a = a_1 < b_1 < \ldots < a_k < b_k = b$ it then follows as above that $M_n(a_1, b_1; \rho), \ldots, M_n(a_k, b_k; \rho)$ have the same distribution as $(1-\rho)^{1/2}M_n(a_1, b_1; 0) + \rho^{1/2}\zeta, \ldots, (1-\rho)^{1/2}M_n(a_k, b_k; 0) + \rho^{1/2}\zeta$ where $\{M_n(a_i, b_i; 0)\}_{i=1}^{k}$ and $\zeta$ all are independent and $\zeta$ is standard normal. Now

$$P\Big(\bigcap_{i=1}^{k}\{N_n((a_i, b_i]) = 0\}\Big) = P\Big(\bigcap_{i=1}^{k}\{M_n(a_i, b_i) \le u_n\}\Big)$$

and with $\rho_n = \gamma/\log n$ it follows from Lemmas 3.2 and 5.1 that

$$P\Big(\bigcap_{i=1}^{k}\{M_n(a_i, b_i) \le u_n\}\Big) - P\Big(\bigcap_{i=1}^{k}\{M_n(a_i, b_i; \rho_n) \le u_n\}\Big) \to 0$$

as $n \to \infty$. Thus, to prove (b) of Theorem A.1 it is enough to check that

$$(5.7) \qquad P\Big(\bigcap_{i=1}^{k}\{M_n(a_i, b_i; \rho_n) \le u_n\}\Big) \to P\Big(\bigcap_{i=1}^{k}\{N((a_i, b_i]) = 0\}\Big).$$

However,

$$P\Big(\bigcap_{i=1}^{k}\{M_n(a_i, b_i; \rho_n) \le u_n\}\Big)$$

$$= P\Big(\bigcap_{i=1}^{k}\{(1-\rho_n)^{1/2}M_n(a_i, b_i; 0) + \rho_n^{1/2}\zeta \le u_n\}\Big)$$

$$= \int_{-\infty}^{\infty} P\Big(\bigcap_{i=1}^{k}\{M_n(a_i, b_i; 0) \le (1-\rho_n)^{-1/2}(u_n - \rho_n^{1/2}z)\}\Big)\phi(z)\,dz,$$

and using that $a_n = \sqrt{2\log n}$, $b_n = a_n + O(a_n^{-1}\log\log n)$, and $\rho_n = \gamma/\log n$ we get

$$(1-\rho_n)^{-1/2}(u_n - \rho_n^{1/2}z)$$

$$= (1 + \rho_n/2 + o(\rho_n))(x/a_n + b_n - \rho_n^{1/2}z)$$

$$= x/a_n + b_n - (\gamma/\log n)^{1/2}z + (\gamma/\log n)\sqrt{2\log n}/2 + o(a_n^{-1})$$

$$= (x + \gamma - \sqrt{2\gamma}z)/a_n + b_n + o(a_n^{-1}).$$

Hence it follows from Theorem 4.1(ii) that, for fixed $z$,

$$P\Big(\bigcap_{i=1}^{k}\{M_n(a_i, b_i; 0) \le (1-\rho_n)^{-1/2}(u_n - \rho_n^{1/2}z)\}\Big)$$

$$\to \prod_{i=1}^{k}\exp\{-(b_i - a_i)e^{-x-\gamma+\sqrt{2\gamma}z}\},$$

and by dominated convergence this proves that

$$\int_{-\infty}^{\infty} P\Big(\bigcap_{i=1}^{k}\{M_n(a_i, b_i; 0) \le (1-\rho_n)^{-1/2}(u_n - \rho^{1/2}z)\}\Big)\phi(z)\,dz$$

$$\to \int_{-\infty}^{\infty}\prod_{i=1}^{k}\exp\{-(b_i - a_i)e^{-x-\gamma+\sqrt{2\gamma}z}\}\phi(z)\,dz$$

$$= P\Big(\bigcap_{i=1}^{k}\{N((a_i, b_i]) = 0\}\Big),$$

i.e. that (5.7) holds. □

<u>COROLLARY 5.3</u> Suppose that the conditions of Theorem 5.2 are satisfied and that $B_1, \ldots, B_k$ are disjoint positive Borel sets whose boundaries have Lebesgue measure zero. Then $P\Big(\bigcap_{i=1}^{k}\{N_n(B_i) = k_i\}\Big)$ tends to the

expression in the right hand side of (5.6). In particular

$$P\{a_n(M_n - b_n) \leq x\} = P\{N_n((0,\ 1]) = 0\} \to \int_{-\infty}^{\infty} \exp(-e^{-x-\gamma+\sqrt{2\gamma}z})\phi(z)\,dz$$

as $n \to \infty$. □

It may be noted that it is quite straightforward to extend the above result to deal with crossings of two or more adjacent levels. However, to avoid repetition we will omit the details of this.

Our next concern is the case when $r_n \log n \to \infty$. Again the problem of exceedances of a fixed level by the dependent sequence can be reduced to considering the exceedances of a random level by an independent sequence, but in this case the random part of the level is "too large"; in the limit the independent sequence will have either no or infinitely many exceedances of the random level. Thus it is not possible to find a normalization that makes the point process of exceedances converge weakly to a non-trivial limit, and accordingly we will only treat the one-dimensional distribution of the maximum. Since the derivation of the general result is complicated we shall consider a rather special case, which brings out the main idea of Mittal and Ylvisaker's proof while avoiding some of the technicalities.

First we note that for $\gamma > 0$, $0 < q < 1$, there is a convex sequence $\{r_n\}_{n=0}^{\infty}$ with $r_0 = 1$, $r_n = \gamma/(\log n)^q$, $n \geq n_o$, for some $n_o \geq 2$. (This is easy to see since $\gamma/(\log n)^q$ is convex for $n \geq 2$, and decreasing to zero.) By Polya's criterion $\{r_n\}$ is a covariance sequence, and we shall now consider a stationary zero mean normal sequence $\{\xi_n\}$ with this particular type of covariance. Further, since $\{r_n\}_{n=0}^{\infty}$ is convex, also $(r_0-r_n)/(1-r_n), \ldots, (r_{n-1}-r_n)/(1-r_n), 0, 0, \ldots$ is convex, so again according to Polya's criterion there is a zero mean normal sequence $\{\zeta_n\}$ with these covariances. Clearly $\xi_1, \ldots, \xi_n$ have the same distribution as $(1-r_n)^{1/2}\zeta_1 + r_n^{1/2}\zeta, \ldots, (1-r_n)^{1/2}\zeta_n + r_n^{1/2}\zeta$, where $\zeta$ is standard normal and independent of $\{\zeta_n\}$.

Putting $M_n' = \max_{1\le k\le n} \zeta_k$ the distribution of $M_n = \max_{1\le k\le n} \xi_k$ therefore is the same as that of $(1-r_n)^{1/2}M_n' + r_n^{1/2}\zeta$. This representation is the key to the proof of Theorem 5.6 below, but before proceeding to use it we shall prove two lemmas. The first one is a "technical" lemma of a type already used twice before.

<u>LEMMA 5.4</u> Let $n' = [n e^{-\sqrt{\log n}}]$ and, suppressing the dependence on n, let $\rho_k = (r_k - r_n)/(1-r_n)$, $k = 1, \ldots, n$. Then, for each $\varepsilon > 0$,

$$n' \sum_{k=1}^{n'} |\rho_k - \rho_{n'}| e^{-(b_n - \varepsilon r_n^{1/2})^2/(1+\rho_k)} \to 0 \text{ as } n \to \infty. \tag{5.8}$$

<u>PROOF</u> Set $p = [n^\alpha]$ where $0 < \alpha < (1-r_1)/(1+r_1)$. Since $r_n$ is decreasing we have as in the proof of Lemma 3.1 that the sum up to $p$ tends to zero, and it only remains to prove that

$$n' \sum_{k=p+1}^{n'} |\rho_k - \rho_{n'}| e^{-(b_n - \varepsilon r_n^{1/2})^2/(1+\rho_k)} \to 0 \text{ as } n \to \infty. \tag{5.9}$$

Now, using (3.4) (with $u_n = b_n$),

$$\begin{aligned} n' \sum_{k=p+1}^{n'} \exp\left\{-\frac{(b_n - \varepsilon r_n^{1/2})^2}{1+\rho_k}\right\} & \\ \sim n e^{-\sqrt{\log n}} \sum_{k=p+1}^{n'} \exp\left\{-\frac{b_n^2}{2}\cdot 2\,\frac{(1-\varepsilon r_n^{1/2}/b_n)^2}{1+\rho_k}\right\} & \\ \sim K n e^{-\sqrt{\log n}} \sum_{k=p+1}^{n'} \left(\frac{b_n}{n}\right)^{2(1-\varepsilon r_n^{1/2}/b_n)^2/(1+\rho_k)} & \\ \le K \log n\, e^{-\sqrt{\log n}} \sum_{k=p+1}^{n'} \frac{1}{n} \exp\left\{2\left(1 - \frac{(1-\varepsilon r_n^{1/2}/b_n)^2}{1+\rho_k}\right)\log n\right\}. & \end{aligned} \tag{5.10}$$

Here $1 - (1-\varepsilon r_n^{1/2}/b_n)^2/(1+\rho_k) \le 1 + \rho_k - 1 + 2\varepsilon r_n^{1/2}/b_n = \rho_k + 2\varepsilon r_n^{1/2}/b_n$, and since

$$r_n^{1/2} b_n^{-1} \log n \le K'(\log n)^{(1-q)/2}$$

we obtain the bound

$$K \log n\, e^{-\sqrt{\log n}} + K'(\log n)^{(1-q)/2} \sum_{k=p+1}^{n'} \frac{1}{n} e^{2\rho_k \log n}. \tag{5.11}$$

Furthermore, for $k \ge p+1$ we have $\log k \ge \alpha \log n$, and then

$$\rho_k \log n = \frac{\gamma/(\log k)^q - \gamma/(\log n)^q}{1-\gamma/(\log n)^q} \log n$$

$$\leq K \frac{\log n}{(\alpha \log n)^q}\left\{1-\left(\frac{\log k}{\log n}\right)^q\right\}$$

$$\leq K (\log n)^{1-q}\left\{1-\left(\frac{\log k/n}{\log n}+1\right)^q\right\}$$

$$\leq K (\log n)^{-q}\left\{-\log k/n\right\},$$

so that

$$\sum_{k=p+1}^{n'} \frac{1}{n} e^{2\rho_k \log n} \leq \sum_{k=p+1}^{n'} \frac{1}{n}\left(e^{-\log k/n}\right)^{K(\log n)^{-q}}$$

$$\leq \int_0^1 \left(e^{-\log x}\right)^{K(\log n)^{-q}} dx$$

$$= \int_0^1 x^{-K(\log n)^{-q}} dx = \frac{1}{1-K(\log n)^{-q}}$$

which tends to one as $n \to \infty$, and therefore is bounded. Together with (5.10) and (5.11) this implies that (5.9) is bounded by

$$K \log n \, e^{-\sqrt{\log n}} + K'(\log n)^{(1-q)/2} \to 0$$

as $n \to \infty$, which proves (5.8). □

<u>LEMMA 5.5</u> For all $\varepsilon > 0$

$$P(|M_n' - b_n| > \varepsilon r_n^{1/2}) \to 0 \quad \text{as} \quad n \to \infty,$$

for $M_n' = \max_{1\leq k\leq n} \zeta_k$, with $\zeta_1, \ldots, \zeta_n$ standard normal and with covariances $\{\rho_k\}$.

<u>PROOF</u> As above write $M_n(\rho)$ for the maximum of $n$ standard normal variables with constant correlation $\rho$ between any two. By definition, $\rho_k \geq 0$, and hence, by (3.6) of Lemma 3.2,

$$P\{M_n' > b_n + \varepsilon r_n^{1/2}\} \leq P\{M_n(0) > b_n + \varepsilon r_n^{1/2} a_n/a_n\}.$$

Further, by definitions, $a_n r_n^{1/2} \to 0$ as $n \to \infty$, and by Theorem 1.11 it follows that

$$P\{M_n' > b_n + \varepsilon r_n^{1/2}\} \to 0.$$

To show that

$$(5.12)\quad P\{M_n' < b_n - \varepsilon r_n^{1/2}\} \to 0$$

we estimate the difference

$$P\{M_n' < b_n - \varepsilon r_n^{1/2}\} - P\{M_n(\rho_{n'}) < b_n - \varepsilon r_n^{1/2}\}$$

by means of (3.6) in Lemma 3.2. Since $\{\rho_k\}$ is convex and thus decreasing, $\rho_k \le \rho_{n'}$ for $k \ge n'$, and we obtain the bound

$$K\, n' \sum_{k=1}^{n'} |\rho_k - \rho_{n'}| e^{-(b_n - \varepsilon r_n^{1/2})^2/(1+\rho_k)}$$

which tends to zero by Lemma 5.4.

Moreover, $M_n(\rho_{n'})$ has the same distribution as $(1-\rho_{n'})^{1/2} M_n(0) + \rho_{n'}^{1/2}\zeta$ where $M_n(0)$ and $\zeta$ are independent, so

$$(5.13)\quad P\{M_n(\rho_{n'}) < b_n - \varepsilon r_n^{1/2}\} = P\{(1-\rho_{n'})^{1/2} M_n(0) + \rho_{n'}^{1/2}\zeta < b_n - \varepsilon r_n^{1/2}\}$$

$$= P\{(1-\rho_{n'})^{1/2} r_n^{-1/2} a_n^{-1} a_n (M_n(0) - b_n) +$$

$$+ ((1-\rho_{n'})^{1/2} - 1) b_n r_n^{-1/2} + \rho_{n'}^{1/2} r_n^{-1/2}\zeta < -\varepsilon\}.$$

For $n$ large,

$$\rho_{n'} = (\gamma/(\log n')^q - \gamma/(\log n)^q)/(1 - \gamma/(\log n)^q)$$

$$\sim \frac{\gamma}{(\log n)^q}\,(1/(1-(\log n)^{-1/2})^q - 1) \sim \gamma q (\log n)^{-1/2-q}.$$

Thus

$$((1-\rho_{n'})^{1/2} - 1)\, b_n r_n^{-1/2} \sim \frac{1}{2}\,\rho_{n'} b_n r_n^{-1/2} \sim q\sqrt{\frac{\gamma}{2}}(\log n)^{-q/2} \to 0$$

as $n \to \infty$ and also $\rho_{n'}^{1/2} r_n^{-1/2} \sim q\,\sqrt{\gamma}(\log n)^{-1/4} \to 0$. Moreover,

$$(1-\rho_{n'})^{1/2} r_n^{-1/2} a_n^{-1} \sim \sqrt{\gamma}(\log n)^{-(1-q)/2} \to 0,$$

and since $a_n(M_n(0) - b_n)$ converges in distribution by Theorem 1.11 it follows that the expression in (5.13) tends to zero, and thus that (5.12) holds. □

<u>THEOREM 5.6</u> Suppose that the stationary, standard normal sequence $\{\xi_n\}$ has covariances $\{r_n\}$ with $\{r_n\}_{n=0}^{\infty}$ convex and $r_n = \gamma/(\log n)^q$, $\gamma \geq 0$, $0 < q < 1$, $n \geq n_0$ for some $n_0$. Then

$$P\{r_n^{-1/2}(M_n-(1-r_n)^{1/2}b_n) \leq x\} \to \Phi(x), \quad \text{as} \quad n \to \infty.$$

<u>PROOF</u> As was noted just before Lemma 5.4, $M_n$ has the same distribution as $(1-r_n)^{1/2}M_n' + r_n^{1/2}\zeta$, where $\zeta$ is standard normal. It now follows at once from Lemma 5.5 that

$$P\{r_n^{-1/2}(M_n - (1-r_n)^{1/2}b_n) \leq x\} = P\{(1-r_n)^{1/2}r_n^{-1/2}(M_n'-b_n) + \zeta \leq x\}$$

$$\to \Phi(x) \quad \text{as} \quad n \to \infty. \qquad \square$$

Of course the hypothesis of Theorem 5.6 is very restrictive. The following more general result was proved by McCormick and Mittal (1976). Their proof follows similar lines as the proof of Theorem 5.6 above, but the arguments are much more complicated.

<u>THEOREM 5.7</u> *Suppose that the stationary standard normal sequence* $\{\xi_n\}$ has covariances $\{r_n\}$ such that $r_n \to 0$ monotonically and $r_n \log n \to \infty$, monotonically for large $n$. Then

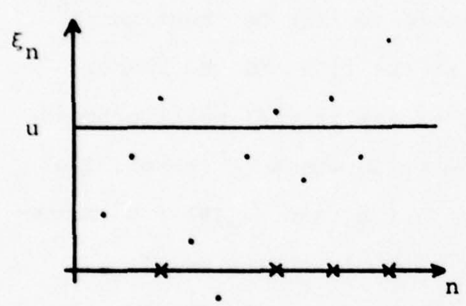
$$P\{r_n^{-1/2}(M_n - (1-r_n)^{1/2}b_n) \leq x\} \to \Phi(x) \quad \text{as} \quad n \to \infty. \qquad \square$$

In the paper by Mittal and Ylvisaker (1975), where the above results were first proved under the extra assumption that $\{r_n\}_{n=0}^{\infty}$ is convex, it is also shown that the limit distributions in Theorems 5.2 and 5.7 are by no means the only possible ones; they exhibit a further class of limit distributions which occur when the covariance decreases irregularly.

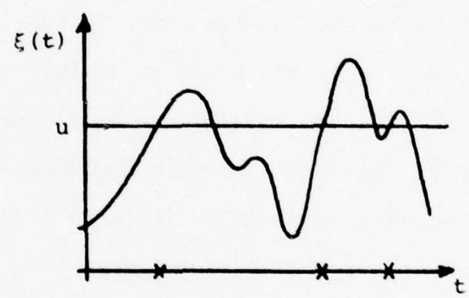
## APPENDIX

## SOME BASIC CONCEPTS OF POINT PROCESS THEORY

Intuitively, by a *point process*, we generally mean a series of events occurring in time (or space, or both) according to some statistical law. For example the events may be radioactive disintegrations or telephone calls, occurring in time, or the positions of a certain variety of plant in a field (two dimensional space). The cases of particular interest to us are when the events are the instants of occurrence of exceedances of a level $u$ by a stochastic sequence $\{\xi_n\}$ (Chapter 4) or of the up-crossings of a level $u$ by a continuous parameter process $\{\xi(t)\}$ (Chapter 8).




Point process of exceedances




Point process of upcrossings

These point processes occur in one dimension (which we may regard as "time" if we wish). We may simultaneously consider exceedances or upcrossings of more than one level, and obtain a point process in the plane (cf. Chapters 4 and 9).

Point process theory may be discussed in a quite abstract setting, leading to a very satisfying general theory, and we refer the interested reader to the books by Kallenberg (1976) and Matthes, Kerstan & Mecke (1978) for this. Here we shall just indicate some of the main concepts regarding point processes on the real line, and in the plane.

If $I$ is any finite interval on the real line, the number of events, $N(I)$ say, of a point process occuring in $I$ must be a random variable.

More generally for any bounded Borel set $B$, $N(B)$ should be a r.v. Further, the number of events in the union of finitely or countably many disjoint sets, is the sum of the numbers in each set, i.e. $N(B) = \sum_1^\infty N(B_i)$ if $B_i$ are disjoint (Borel) sets whose union is $B$. That is $N(\cdot)$ is a *measure* on the Borel sets. Again the value of $N(B)$ must be an integer as long as it is finite. Hence the following formal definition naturally suggests itself.

*A point process in* $R^n$ *is a family of non-negative integer (or* $+\infty$*)-valued r.v.'s* $N_\omega(B)$ *defined for each Borel set* $B$ *and such that for each* $\omega$, $N_\omega(\cdot)$ *is a measure on the Borel sets, being finite valued on bounded sets.*

Note that the same definition may be used for more abstract topological spaces by using the Borel set of that space in lieu of those in $R^n$. Here, as noted, we shall concider mainly just the line and the plane.

If $\tau$ is a random variable we may consider the trivial point process consisting of just one event occurring at the value which $\tau$ takes, i.e. the point process $\delta_\tau$ where $\delta_\tau(B) = 1$ if $\tau \in B$ and $\delta_\tau(B) = 0$ otherwise. Thus $\delta_\tau$ represents unit mass at the randomly chosen point $\tau$. More generally if $\tau_j$ are random variables (for $j = 1, 2, 3, \ldots,$ or $j = 0, \pm 1 \pm 2, \ldots$) we may define a point process $N = \sum_j \delta_{\tau_j}$ provided that $N(B) = \sum_j \delta_{\tau_j}(B)$ is finite a.s. for bounded sets $B$. For this point process, the events occur at the random points $\{\tau_j\}$. Going a little further still, we may conveniently include possible *multiple events* by writing $N = \sum_j \beta_j \delta_{\tau_j}$ where $\beta_j$ are non-negative integer valued r.v.'s, and the $\tau_j$ taken distinct.

In fact for a space with sufficient structure (such as the real line or plane) it may be shown that any point process $N$ may be represented in terms of its *atoms* in this way, i.e. $N = \sum_j \beta_j \delta_{\tau_j}$, where the $\tau_j$ are a.s. distinct random elements of the space, and $\beta_j$ are non-negative, integer valued random variables. In the case where the $\beta_j$ are each unity a.s., we say that the point process *has no multiple events* or is *simple*.

If $B_1, \ldots, B_k$ are bounded Borel sets, $N(B_1), \ldots, N(B_k)$ are random variables and have a joint distribution — termed a *finite dimensional distribution* of the point process. In fact, the probabilistic properties of interest concerning the point process are specified uniquely by the collection of all such finite dimensional distributions, i.e. for all choices of $k$ and the sets $B_1, \ldots, B_k$. Of course, to define a point process starting from finite dimensional distributions, we must choose these distributions in an appropriately consistent manner, so that the r.v.'s $N(B)$ will not only be well defined but will be non-negative, countably additive in $B$, etc. We refer the interested reader to Kallenberg (1976) for details on this.

If $N$ is a point process the measure $\lambda$ defined on the Borel sets (of the space involved) by

$$\lambda(B) = E(N(B))$$

is termed the *intensity measure* of the point process. Note that, unlike $N(B)$ itself, $\lambda(B)$ may be infinite even when $B$ is bounded (since while a r.v. is finite valued, its mean need not be).

Although we shall not need them here, it is of interest to note that the probabilistic properties of a point process $N$ may also be summarized by various generating functionals. In our judgement the most natural and useful of these is the Laplace transform $L_N(f)$ defined for non-negative measurable functions $f$ by

$$L_N(f) = E(e^{-\int f dN}) = E(e^{-\Sigma \beta_j f(\tau_j)})$$

when $N$ is represented as $\Sigma \beta_j \delta_{\tau_j}$. Such generating functionals have properties and uses analogous to those of characteristic function, moment generating functions, and Laplace transforms of random variables. In particular if $f(x) = t\,\chi_B(x)$ (where $\chi_B(x) = 1$ or $0$ according as $x \in B$ or $x \notin B$) we have $L_N(f) = E(e^{-tN(B)})$. This is simply the Laplace transform (or moment generating function evaluated at $-t$) of the r.v. $N(B)$, and it uniquely specifies the distribution of $N(B)$.

AD-A069 561 NORTH CAROLINA UNIV AT CHAPEL HILL INST OF STATISTICS F/G 12/1
EXTREMAL AND RELATED PROPERTIES OF STATIONARY PROCESSES. PART I--ETC(U)
MAY 79 M R LEADBETTER, G LINDGREN, H ROOTZEN N00014-75-C-0809

UNCLASSIFIED MIMEO-SER-1227 NL

2 OF 2
AD
A069561

END
DATE
FILMED
7-79
DDC

Similarly joint Laplace transforms for r.v.'s $N(B_1), \ldots, N(B_k)$ may be specified by taking $f = \Sigma_{i=1}^{k} t_i \chi_{B_i}$.

Probably the most useful point process – both in its own right, and also as a "building block" for other types – is the *Poisson process*. This may be specified by its intensity measure $\lambda(B)$ which may be taken to be any measure which is finite on bounded sets. The point process $N$ is said to be Poisson with this intensity if for each (bounded) $B$, $N(B)$ is a Poisson r.v. with mean $\lambda(B)$, and $N(B_1), \ldots, N(B_k)$ are independent for any choice of $k$, and disjoint $B_1, \ldots, B_k$. The existence of such a process $N$ is easily shown under very general circumstances though we do not do so here, and $N$ has the Laplace transform $L_N(f) = \exp\{-\int(1 - e^{-f(u)})d\lambda(u)\}$.

It is readily checked that the choice $f(x) = t\chi_B(x)$ yields the Laplace transform of a Poisson r.v. with mean $\lambda(B)$. Incidentally this process may be called the "general Poisson process". The usual (stationary) Poisson process on the real line arises when $\lambda(B)$ is a constant multiple of Lebesgue measure $m(B)$, i.e. $\lambda(B) = \tau m(B)$, in which case we say that the Poisson process har intensity $\tau$.

As noted above, the Poisson process may be used as a building block for the construction of other point processes. In particular, a most useful case arises if the intensity measure $\lambda$ is itself allowed to be stochastic. Such a point process is no longer Poisson, but may be profitably thought of as "Poisson with a (stochastically) varying mean rate". We refer to such a process as a *Doubly stochastic Poisson* or, more commonly, a *Cox process*. Specific use is made of such processes in Chapter 5, where the distribution is explicitly given.

A notion which will be useful to us is that of *thinning* of a point process – and, in particular of a Poisson process. Thinning refers to the removal of some of the events of the point process by a (usually) probabilistic mechanism which can be quite complicated. In its simplest form – with which we shall be concerned here – each event is removed or

retained *independently*, with probabilities $1-p, p$ say. For example if $N$ is a Poisson process with intensity measure $\lambda$, and $N^*$ a point process obtained from $N$ by such independent thinning, we have, for a Borel set $B$,

$$P\{N^*(B)=r\} = \sum_{s=r}^{\infty} P\{N(B)=s\} P\{N^*(B)=r \mid N(B)=s\}$$

$$= \sum_{s=r}^{\infty} \frac{e^{-\lambda(B)}(\lambda(B))^s}{s!} \binom{s}{r} p^r (1-p)^{s-r},$$

since given that $N(B)=s$, $N^*(B)$ is binomial with parameters $(s, p)$. This expression reduces simply to yield

$$P\{N^*(B)=r\} = e^{-p\lambda(B)}(p\lambda(B))^r/r!$$

so that $N^*(B)$ is a Poisson r.v. with mean $p\lambda(B)$. Similarly it may be seen that $N^*(B_1), \ldots, N^*(B_k)$ are independent whenever $B_1, \ldots, B_k$ are disjoint, so that $N^*$ is clearly a Poisson process with intensity measure $p\lambda$ – an intuitively appealing result of which use is made e.g. in Chapter 4.

There are numerous structural properties of point process theory – such as existence, uniqueness, simplicity, infinite divisibility and so on – which we do not go into here. It will however be of interest to mention *convergence* of a sequence of point processes, and to state a useful theorem in this connection.

Suppose that $\{N_n\}$ is a sequence of point processes and that $N$ is a point process. Then we may say that $N_n$ *converges in distribution* to $N$ (written $N_n \xrightarrow{d} N$) if the sequence of vector r.v.'s $(N_n(B_1), \ldots, N_n(B_k))$ converges in distribution to $(N(B_1), \ldots, N(B_k))$ for each choice of $k$, and all bounded Borel sets $B_i$ such that $N(\partial B_i)=0$ a.s., $i=1, \ldots, k$, (writing $\partial B$ for the bounary of the set $B$).

A point process may be viewed as a random element of a certain metric space (whose points are measures) and convergence in distribution of $N_n$ to $N$ becomes weak convergence of the distributions of $N_n$ to that of $N$. However we do not need this general viewpoint here since the above

definition is equivalent to it.

The main result which we shall need is the following simple sufficient condition for convergence in distribution. This is a special case of a theorem of Kallenberg (1976) and we state it here without proof.

<u>THEOREM A.1</u> (i) Let $N_n$, $n = 1, 2, \ldots$, and $N$ be point processes on the real line, $N$ being simple. Suppose that

(a) $E(N_n((a, b])) \to E(N((a, b]))$ for all $-\infty < a < b < \infty$

and

(b) $P\{N_n(B) = 0\} \to P\{N(B) = 0\}$ for all $B$ of the form $\bigcup_1^k (a_i, b_i]$

$k = 1, 2, \ldots$

Then $N_n \xrightarrow{d} N$.

(ii) The same is true for point processes in the plane, if the semiclosed intervals $(a, b]$, $(a_i, b_i]$ are replaced by "semiclosed rectangles" $(a, b] \times (\alpha, \beta]$, $(a_i, b_i] \times (\alpha_i, \beta_i]$. □

The remarkable feature of this result if that convergence of the probability of occurrence of no events in certain given sets is essentially sufficient to guarantee convergence of quantities like $P\{N_n(B) = r\}$ and corresponding joint probabilities. The simple conditions (a) and (b) are often readily verified.

Our next result concerns convergence of a sequence of point processes in the plane, to a Poisson process in the plane, and shows how this property is preserved under suitable transformations of the points of each member of the sequence, and of the limit. Obviously this result could be stated in much greater generality but the form given here is sufficient for our applications.

<u>THEOREM A.2</u> Let $N_n$, $n = 1, 2, \ldots$, and $N$ be point processes in the plane, and $\iota(x)$ a strictly decreasing continuous real function. Define new point processes $\{N_n'\}$, $N'$ such that if $N_n(N)$ has an atom at $(s, t)$

then $N_n'(N')$ has an atom at $(s, \tau^{-1}(t))$, where $\tau^{-1}$ is the inverse function of $\tau$.

(i) If $N_n \xrightarrow{d} N$ then $N_n' \xrightarrow{d} N'$.

(ii) If $N$ is Poisson, with intensity measure $\lambda$, then $N'$ is Poisson with intensity $\lambda T^{-1}$ where $T$ denotes the transformation of the plane given by $T(s, t) = (s, \tau^{-1}(t))$. If $\lambda$ is Lebesgue measure on the plane, the intensity $\lambda T^{-1}$ is the product of linear Lebesgue measure and the measure defined by the monotone function $\tau$.

PROOF (i) It is readily checked that for any rectangle $B = (a, b] \times (\alpha, \beta]$

$$N'(B) = N((a, b] \times [\tau(\beta), \tau(\alpha)]) = N(T^{-1}(B))$$

and hence (by uniqueness of extensions of measures) $N'(B) = N(T^{-1}(B))$ for all Borel sets $B$. This holds also with $N_n'$, $N$ replacing $N'$, $N$.

Suppose now that $B$ is a Borel set such that $N'(\partial B) = 0$ a.s. (again $\partial B$ denotes the boundary of $B$). Now it may be seen (using the continuity of $\tau$) that $\partial T^{-1} B \subset T^{-1} \partial B$ so that $N(\partial T^{-1} B) \leq N(T^{-1} \partial B) = N'(\partial B) = 0$ a.s. Since $N_n \xrightarrow{d} N$ we thus have $N_n(T^{-1}B) \xrightarrow{d} N(T^{-1}B)$ or $N_n'(B) \xrightarrow{d} N'(B)$. This result extends simply to show that $(N_n'(B_1), \ldots, N_n'(B_k)) \xrightarrow{d} (N'(B_1), \ldots, N'(B_k))$ whenever $N'(\partial B_i) = 0$ a.s. for each $i = 1, \ldots, k$ and hence $N_n' \xrightarrow{d} N'$ as required.

(ii) If $N$ is Poisson with intensity $\lambda$, and $B$ is any Borel set in the plane,

$$P\{N'(B) = r\} = P\{N(T^{-1}(B)) = r\} = e^{-\lambda(T^{-1}B)} (\lambda(T^{-1}B))^r / r!$$

for each $r = 0, 1, 2, \ldots$, so that $N'(B)$ is Poisson with mean $\lambda T^{-1}(B)$. Independence of $N'(B_1), \ldots, N'(B_k)$ for disjoint $B_1, \ldots, B_k$ follows from the fact that $T^{-1}B_1, \ldots, T^{-1}B_k$ are also disjoint, and hence $N(T^{-1}B_1), \ldots, N(T^{-1}B_k)$ are independent.

The last statement of (ii) follows simply since if $\lambda$ is Lebesgue measure,

$$\lambda T^{-1}((a, b]\times(\alpha, \beta]) = \lambda\{(a, b]\times[\tau(\beta), \tau(\alpha))\}$$
$$= (b-a)(\tau(\alpha)-\tau(\beta)),$$

noting also that $\tau$ is continuous. □

As a final note, it is apparent that the concept of a point process may be generalized to include measures $N$ for which $N(B)$ is not necessarily integer valued. This generalization leads to a natural setting for point processes within the framework of the theory of Random Measures – a viewpoint developed in detail by Kallenberg (1976).

## REFERENCES

Adler, R.J. (1978). Weak convergence results for extremal processes generated by dependent random variables. To appear.

Berman, S.M. (1964). Limit theorems for the maximum term in stationary sequences. Ann. Math. Statist. 35, 502-516.

Berman, S.M. (1971 a). Asymptotic independence of the numbers of high and low level crossings of stationary Gaussian processes. Ann. Math. Statist. 42, 927-947.

Chibisov, D.M. (1964). On limit distributions of order statistics. Prob. Appl. 9, 142-148.

Cramér, H. (1965). A limit theorem for the maximum values of certain stochastic processes. Teor. Ver. i Prim. 10, 137-139.

Cramér, H. & Leadbetter, M.R. (1967). Stationary and Related Stochastic Processes, New York: John Wiley and Sons.

Davis, R.A. (1978). Maxima and minima of stationary sequences. To appear.

Fisher, R.A. & Tippet, L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc. Cambridge Phil. Soc. 24, 180-190.

Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. Ann. Math. 44, 423-453.

Gumbel, E.J. (1958). Statistics of Extremes. New York: Columbia Univ. Press.

Haan, L. de (1970). On regular variation and its application to the weak convergence of sample extremes. Amsterdam Math. Centre Tract 32.

Haan, L. de (1976). Sample extremes: an elementary introduction. Statistica Neerlandica 30, 161-172.

Kallenberg, O. (1975). Random measures. Berlin: Akad. Ver.

Leadbetter, M.R. (1974). On extreme values in stationary sequences. Z. Wahr. verw. Geb. 28, 289-303.

Leadbetter, M.R., Lindgren, G. & Rootzén, H. (1979). Conditions for the convergence in distribution of maxima of stationary normal processes. Stochastic Processes Appl. 8.

Loynes, R.M. (1965). Extreme values in uniformly mixing stationary stochastic processes. Ann. Math. Statist. 36, 993-999.

McCormick & Mittal (1976). On weak convergence of the maximum. Techn. Rep. No. 81, Stanford University.

Mittal, Y. & Ylvisaker, D. (1975). Limit distributions for maxima of stationary Gaussian processes. Stochastic Processes Appl. 3, 1-18.

Pickands, J. (1971). The two-dimensional Poisson process and extremal processes. J. Appl. Prob. 8, 745-756.

Resnick, S. (1975). Weak convergence of extremal processes. Ann. Prob. 3, 951-960.

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. Proc. Nat. Acad. Sci. USA 42, 43-47.

Slepian, D. (1962). *The one-sided* barrier problem for Gaussian noise. Bell Syst. Tech. J. 41, 463-501.

Smirnov, N.V. (1952). The limit distributions for the terms of a variational series. Ann. Math. Soc. Translation 67.

Smirnov, N.V. (1967). Some remarks on limit laws for order statistics. T. Prob. Appl. 12, 336-337.

Watson, G.S. (1954). Extreme values in samples from m-dependent stationary stochastic processes. Ann. Math. Statist. 25, 798-800.

Wu, Chuan-Yi (1966). The types of limit distributions for some terms of variational series. Scientia Sinica 15, 749-762.